Efficient Demands in Multi-product Monopoly*

Kai Hao Yang†

April 18, 2021

Abstract

This paper characterizes the efficient market demands among those with a fixed surplus level in a multi-product monopoly where the monopolist is able to produce a continuum of quality-differentiated products with a cost function that is convex in quality. We show that any efficient market demand must be *affine-unit-elastic*. This further reduces the problem of characterizing the efficient frontier to a finite-dimensional constraint optimization problem. From this characterization, it follows that deadweight losses are positive even under efficient demands; that both consumer surplus and monopoly profit are nonmonotonic in cost function; and that the monopolist sells at most two distinct quality levels under any efficient market demand.

Keywords: Multi-product monopoly, second-degree price discrimination, upgrade markets, efficient demand, affine-unit-elastic demand, technology.

Jel classification: D42, D82, L11, L12, L15

*This paper was previously circulated under the title "Strategic Information Acquisition in Auctions", "Buyer-Optimal Information under Uncertainty", and "Buyer-Optimal Information with Nonlinear Technology" This version supersedes the previous ones. It also subsumes part of the results of a companion paper "Informationally Robust Welfare Predictions under Second-Degree Price Discrimination".

†Cowles Foundation for Economic Research, Yale University, kaihao.yang@yale.edu. I am indebted to Ben Brooks, Elliot Lipnowski, and Doron Ravid for their guidance during the early stage of this project, which establishes the foundation of the main result. I am grateful for the discussions with and comments from Simon Board, Steven Davis, Alex Frankel, Emir Kamenica, Eric Maskin, Stephen Morris, Roger Myerson, Phil Reny, and Nancy Stokey. I especially thank the associate editor and three anonymous referees for their valuable suggestions that lead to this version.

1 Introduction

For more than a century, monopoly pricing has been one of the most salient part of modern economic analysis. Various streams of literature and numerous studies have widely explored the welfare consequences and the allocative outcomes in monopoly markets. Naturally, the implications on welfare and allocation largely depend on features of the market demand and the technology. In other words, a comprehensive understanding of the welfare and allocative effect of market demands and technology is essential for the theory of monopoly. After all, market demands and technologies are by no means constant across time and space, and can be greatly influenced by various economic activities. For instance, the demand function in a market can be affected by product design, marketing, advertisement, consumers' income distribution, or change of conditions in other markets; whereas the production technology can be altered by innovation, product line management, or vertical integration. Consequently, a natural question arises: How do market demands and technologies affect the welfare and allocative outcomes in a monopoly market?

In a seminal paper, Johnson and Myatt (2006b) approach this question from a perspective of comparative statics. They consider a particular (one-dimensional) form of changes in demands (i.e., the *rotational order*, which measures how dispersed the consumers values are) that can naturally arise under canonical models of product design, advertisement, and product line transformations. Within this one-dimensional demand changes, they show that in both a single-product monopoly and a quality-differentiated multi-product monopoly, the monopolist's profits are quasi convex, meaning that the monopolist would prefer the "least dispersed" or the "most dispersed" market demands.

Parallel to the one-dimensional comparative statics analysis, Roesler and Szentes (2017) approach the same question by focusing on a single-product monopoly setting, and exploring the entire set of welfare outcomes that can be induced by an infinite dimensional set market demands (i.e, those satisfying a mean preserving contraction constraint).[1] While restricting attention to a single-product monopoly brings a great amount of tractability and allows them to fully characterize the entire set of monopoly profit and consumer surplus, as well as the demand functions that induce each of the payoff pairs, it rules out the possibility that the monopolist can produce quality-differentiated products. This exclusion is crucial since a monopolist would generally offer different products and engage in second-degree price discrimination when she can produce quality-differentiated products (see, for instance, Mussa and Rosen (1978)), rather than selling a single product alone.

[1]Although motivated as buyer's information, a marginal distribution of interim expected value in Roesler and Szentes (2017) is in fact equivalent to a market demand satisfying a mean preserving contraction constraint. Relatedly, some of the results of Condorelli and Szentes (2020) can also be regarded as characterizing the consumer-optimal demands under various constraints induced by certain information costs.

In this paper, we take the infinite dimensional approach of Roesler and Szentes (2017) and explore the welfare and allocative consequences of market demands in a multi-product monopoly. Specifically, consider a monopolist who is able to produce a continuum of quality-differentiated products with a cost function that is convex in quality level. The main result of this paper characterizes the efficient market demands among those with the same surplus level (with a fixed highest possible consumer value).[2] That is, we solve for the market demands under which no other demands with the same level of total surplus can induce higher consumer surplus and monopoly profit at the same time. We show that for any given surplus level and for any (well-behaved) technology, any efficient market demand must be *affine-unit-elastic*, in the sense that it is an affine transformation of a (truncated) unit-elastic demand. This effectively reduces the problem of characterizing the efficient frontier to a finite-dimensional problem, even though the set of feasible demands is infinite-dimensional.

Using the main characterization, we are able to derive further implications regarding the welfare and the allocative outcomes, as well as the effect of technologies. Specifically, we show that deadweight losses are generically positive even under the efficient market demands. This is in contrast to the case of a single-product monopoly, where the results of Roesler and Szentes (2017) (surprisingly) imply that the deadweight loss must be zero under all efficient market demands. Furthermore, the characterization of the efficient frontier allows us to conduct comparative statics analyses regarding the effect of production technologies. Specifically, we show that the optimal consumer surplus is quasi-convex in technologies ranked by the rotational order; and that the efficient payoff frontier may shift "inward" even when production costs for each quality level decrease. These suggest that the effect of changes in production technology is generally nonmonotonic in a multi-product monopoly. Lastly, as a feature of affine-unit-elastic demands, the main characterization implies that only two quality levels are sold under any efficient market demand. This provides a possible rationale for the use of simple menus, and indicates that the market outcome can be improved via demand manipulation when a monopolist is observed to sell numerous quality-differentiated products.

Compared with a single-product monopoly, the optimal selling mechanism in a multi-product monopoly is much more complicated. Indeed, instead of charging a one-dimensional optimal monopoly price, the monopolist would engage in second-degree price discrimination and use a menu to screen the consumers. As a result, the arguments of Roesler and Szentes

[2] When defining efficiency, we allow for all possible demands with a fixed surplus level. The reason is that, on one hand, after characterizing the efficient demands for each fixed surplus level, we can then obtain the entire efficient frontier among every possible demands by varying the one-dimensional surplus level. On the other hand, sometimes there may be an exogenous reasons that prevent the demands from changing in level and only allow for changes in curvatures (see Johnson and Myatt (2006b)). We discuss more about this assumption in Section 6.

(2017) cannot be extended to a multi-product monopoly setting, as they heavily rely on the simple structure of a single-product monopoly pricing problem. To overcome this technical challenge, we first follow the insight of Johnson and Myatt (2003), Johnson and Myatt (2006a), Johnson and Myatt (2006b) and transform a multi-product monopoly model into a continuum of single-product monopoly pricing problems where the monopolist operates in different "upgrade markets" under different (constant) marginal costs. This allows us to rewrite the monopolist's profit and consumer surplus in a multi-product monopoly into a mixture of single-product monopoly profit and consumer surplus over a continuum of upgrade markets, with weights described by the inverse marginal cost function. Then, we adopt a local perturbation argument and show that, any market demand that is not affine unit elastic can always be locally perturbed and transformed into another marker demand that (i) has the same surplus level, (ii) induces approximately the same monopoly profit in every upgrade market, and (iii) induces higher consumer surplus in a positive measure of upgrade markets. This then implies that any efficient market demand must be affine-unit-elastic. Finally, existence of efficient market demands are ensured by a technical result ensuring that the consumer surplus and the monopoly profit are upper-semicontinuous and continuous in market demand, respectively

The rest of this paper is organized as follows. In the next section, we discuss the related literature. In Section 3, the model and the definition of efficient market demands are introduced. The main result, as well as the outline of its proof is stated in Section 4, followed by further implications in Section 5. Lastly, Section 6 discusses the assumptions, interpretations, and implications of the model while Section 7 conclude.

2 Related Literature

As discussed in Section 1, this paper is closely related to Johnson and Myatt (2006b), Roesler and Szentes (2017). Specifically, while Johnson and Myatt (2006b) consider comparative statics regarding a parameterized change in demand, this paper characterizes the efficient ones among those with the same surplus level. Meanwhile, while Roesler and Szentes (2017) solve for the efficient market demands among those satisfying an arbitrary mean preserving contraction constraint in a single-product monopoly, this paper considers a similar problem in a multi-product monopoly and relaxes the mean preserving contraction constraint to a surplus level constraint.[3] Relatedly, Condorelli and Szentes (2020) solves for the consumer-optimal demand in a single-product monopoly setting where the choice of different demands costs differently, whereas the choice of demand is costless (as long as they have the same surplus

[3]Equivalently, using their interpretation, this paper focuses on a prior with binary support $\{0, 1\}$ whereas Roesler and Szentes (2017) allows for general prior

level) in this paper.[4] It is noteworthy that unit elastic demands are crucial for both Roesler and Szentes (2017) and Condorelli and Szentes (2020). In the meantime, Haghpanah and Siegel (2020) and Haghpanah and Siegel (2021) also study welfare consequences of demands in a multi-product monopoly. They focus on how market segmentations—ways to split the market demand into several market segments to facilitate price discrimination—affect monopoly profit and consumer surplus.

Unit-elastic demands and their variants have also appeared in many other papers in various literatures. A crucial property of this family of demands is that a (single-product) monopoly is indifferent among all the prices at which the demand function is strictly decreasing, which in turn is related to the worst-case demands in robust pricing or auction problems (e.g., Neeman (2003), Bergemann and Schlag (2011), Brooks (2013), Libgober and Mu (forthcoming)), supporting an equilibrium strategy (e.g., Varian (1980), Renou and Schlag (2010), Ravid, Roseler, and Szentes (2019)), or enhancing consumer surplus (e.g., Condorelli and Szentes (2021), Bergemann, Brooks, and Morris (2015)).

As explained by Johnson and Myatt (2006b), many economic activities can shape the market demand. Therefore, this paper is also related to Lancaster (1975), who considers a product design problem that may change the dispersion of consumers' values due to taste. Meanwhile, information is also a factor that would affect demand functions. From this perspective, this paper is related to the advertisement literature and the recent development of information design. Specifically, Lewis and Sappington (1991) studies an advertisement model where truth-or-noise information is given to the consumers and show that the monopolist would prefer the consumers to be either fully informed or completely uninformed. Libgober and Mu (forthcoming) studies a dynamic robust monopoly pricing problem where the monopolist evaluates the buyer's information by a worst-case criterion. Bergemann and Pesendorfer (2007) examine an optimal auction problem where the seller can disclose any information to (uninformed) buyers independently. Methodologically, the model of this paper is equivalent to an information design problem where only the expectation of that state variable is payoff relevant. However, since consumer surplus is a non-convex function of market demands, the duality method of Dworczak and Martini (2019) and the extreme-point method of Kleiner, Moldovanu, and Strack (forthcoming) do not apply.

[4]In fact, out surplus constraint corresponds to the special case of *mean-based costs* in Condorelli and Szentes (2020).

3 Model

3.1 Primitives

A monopolist (she) sells a continuum of quality-differentiated products to a unit mass of consumers. Quality levels are indexed by $q \in [0,1]$. Consumers have unit demand (i.e., each consumer buys at most one quality level) and quasi-linear utility with heterogeneous values distributed on $[0,1]$. For a consumer with value v, his utility from buying quality q and paying p is $vq - p$.

3.2 Technology

To produce a product with quality q, the monopolist incurs cost $C(q) \geq 0$. We focus on the case where the cost function C is convex and does not have a fixed cost (i.e., $C(0) = 0$). For any convex function $C : \mathbb{R}_+ \to \mathbb{R}_+$, define the marginal cost function $C' : \mathbb{R}_+ \to \mathbb{R}_+$ as the right derivative[5] of C and define $\gamma : \mathbb{R}_+ \to [0,1]$ as

$$\gamma(c) := \inf\{q \in [0,1] | C'(q) \geq c\}, \, \forall c. \tag{1}$$

By definition, $\gamma : \mathbb{R}_+ \to [0,1]$ is nondecreasing and right-continuous. As a result, every convex cost function C without a fixed cost induces a nondereasing and right-continuous function $\gamma : \mathbb{R}_+ \to [0,1]$. Conversely, any nondecreasing and right-continuous function $\tilde{\gamma} : \mathbb{R}_+ \to [0,1]$ uniquely defines a convex cost function $\tilde{C}$ without a fixed cost.[6] Henceforth, we use a nondecreasing and right continuous function $\gamma : \mathbb{R}_+ \to [0,1]$ to describe the technology, and denote the set of these functions as Γ. Moreover, we say that a technology $\gamma \in \Gamma$ is *regular* if there exists $\bar{c} \geq 1$ such that γ is strictly increasing and Lipschitz continuous on $[0, \bar{c}]$.[7]

---

[5] That is,

$$C'(q) := \lim_{q' \downarrow q} \frac{C(q') - C(q)}{q' - q}, \, \forall q \in [0,1]$$

[6] To see this, let $\tilde{c}(q) := \inf\{c \geq 0 | \tilde{\gamma}(q) \geq c\}$ for all $q \in [0,1]$ and let $\tilde{C}(q) := \int_0^q \tilde{c}(z) \, \mathrm{d}z$.

[7] This means that every regular γ corresponds to a cost function that is strictly convex cost with $C(0) = 0$ and $C(1) \geq 1$. The assumption that $C(1) \geq 1$ is not essential but it simplifies the notation since otherwise market demands would have a payoff-irrelevant part (i.e., when $v \geq \bar{c}$), which makes the characterization less sharp. Meanwhile, the continuity and monotonicity assumptions are crucial since the proof of the main result relies on a local perturbation argument.

3.3 Market Demand

The distribution of consumers' values v is described by a *market demand* D, where $D : \mathbb{R}_+ \to [0, 1]$ is a nonincreasing and upper-semicontinuous[8] function with $D(0) = 1$ and $D(1^+) = 0$. The interpretation is that for any price p, the share of consumers who are willing to buy the product with quality q (i.e., the share of consumers with $v \geq p/q$) would be $D(p/q)$.

Given a market demand D, the amount of total surplus in the economy is given by the area below the demand curve:[9]

$$\int_0^1 D(v)\,\mathrm{d}v.$$

As discussed in Section 1, we explore every possible market demands that yield the same level of total surplus. As such, for any $s \in (0, 1)$ let

$$\mathcal{D}_s := \left\{ D : \mathbb{R}_+ \to [0,1] \middle| D \text{ is a market demand, } \int_0^1 D(v)\,\mathrm{d}v = s \right\}$$

be the collection of market demands that have surplus level s.

3.4 Market Outcomes

Given any market demand D and any technology γ, as shown by Johnson and Myatt (2003), Johnson and Myatt (2006a), and Johnson and Myatt (2006b), the optimal selling mechanism for a multi-product monopolist (Mussa and Rosen, 1978) is equivalent to a collection of single-product monopoly pricing problems. That is, we may describe the market outcomes by considering the monopolist selling in a continuum of "upgrade markets" at different (constant) marginal cost for each upgrade. The monopoly profit (consumer surplus, resp.) is given by the mixture of profits (surplus, resp.) of each individual upgrade market, according to the (nondecreasing and right continuous) function γ.[10]

[8]Upper-semicontinuity is equivalent to consumers breaking tie in favor of the monopolist, which is necessary for the existence of the monopolist's optimal selling mechanism. This can be guaranteed if we view the model as an extensive form game where the monopolist set prices first and consumers make purchase decisions afterward.

[9]Equivalently, total surplus equals to the expected value under the probability measure μ^D associated with the demand D. Using integration by parts, we have

$$\int_0^1 v\mu^D(\mathrm{d}v) = -vD(v)\Big|_{v=0}^{v=1} + \int_0^1 D(v)\,\mathrm{d}v = \int_0^1 D(v)\,\mathrm{d}v.$$

[10]Although Johnson and Myatt (2003), Johnson and Myatt (2006a), and Johnson and Myatt (2006b) derive this equivalence under the assumption that there are only finitely many quality levels, the result can be easily extended to the model with a continuum of quality levels. A formal statement and its proof can be found in Lemma 2 of Yang (2020),

More specifically, the monopolist's optimal profit under market demand D and technology γ is a mixture of single-product monopoly profits across every upgrade markets:

$$\Pi(D|\gamma) := \int_0^{\bar{c}} \max_{p \geq 0}[(p-c)D(p)]\gamma(\mathrm{d}c),$$

whereas the consumer surplus is a mixture of their surplus across every upgrade markets:

$$\Sigma(D|\gamma) := \int_0^{\bar{c}} \left(\int_{\boldsymbol{p}_D(c)}^1 D(v)\,\mathrm{d}v \right) \gamma(\mathrm{d}c),$$

where $\boldsymbol{p}_D(c)$ is the smallest element of $\operatorname{argmax}_{p \geq 0}(p-c)D(p)$.[11] Henceforth, for each $c \geq 0$, we refer the single-product monopoly pricing problem with marginal cost $c \geq 0$ as *upgrade market c*.

4 Efficient Market Demands

In this section, we present the main result of this paper. As described in Section 1, our main goal is to characterize the efficient market demands for any fixed surplus level $s \in (0,1)$ and any regular technology $\gamma \in \Gamma$. Efficient market demands are those that yield undominated pairs of monopoly profit and consumer surplus. Specifically:

Definition 1. For any $s \in (0,1)$ and for any $\gamma \in \Gamma$, a market demand D is (s,γ)-*efficient* if:

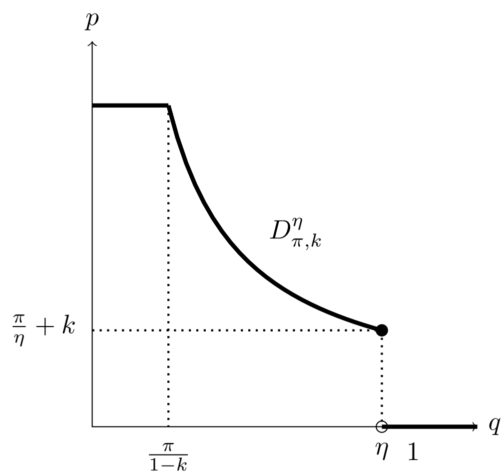
1. $D \in \mathcal{D}_s$;
2. There does not exist any $D' \in \mathcal{D}_s$ such that $\Pi(D'|\gamma) \geq \Pi(D|\gamma)$ and $\Sigma(D'|\gamma) \geq \Sigma(D|\gamma)$, with at least one inequality being strict.

To state the main results, we first introduce a crucial class of market demands, which we call the *affine-unit-elastic demands*. For any $\pi, k, \eta \geq 0$ such that $\pi/\eta + k \leq 1$, define the market demand $D^{\eta}_{\pi,k}$ as follows:

$$D^{\eta}_{\pi,k}(v) := \begin{cases} 1, & \text{if } v = 0 \\ \eta, & \text{if } v \in \left(0, \frac{\pi}{\eta} + k\right] \\ \frac{\pi}{v-k}, & \text{if } v \in \left(\frac{\pi}{\eta} + k, 1\right] \\ 0, & \text{if } v > 1 \end{cases},$$

[11] Here, selecting the smallest optimal price is without loss because our main interest is in maximizing welfare. To see this, notice that whenever there are multiple optimal prices for the monopolist, consumers would always be weakly better-off if the monopolist selects a lower optimal price, while the monopolist would always be indifferent among all optimal prices.

Figure 1: Affine-Unit-Elastic Demand $D^{\eta}_{\pi,k}$



for all $v \geq 0$.[12]

Figure 1 plots (the inverse demand of) $D^{\eta}_{\pi,k}$.[13] Notice that $D^{\eta}_{\pi,k}$ has two jumps, one at $v = 0$ with size $1 - \eta$ and the other at $v = 1$ with size $\pi/(1 - k)$ (and hence the inverse demand in Figure 1 has two flat regions, $[0, \pi/(1 - k)]$ and $[\eta, 1]$). Furthermore, the affine transformation $D^{\eta}_{\pi,k}(v + k)$ has unit elasticity on the interval $(\pi/\eta, 1 - k]$. In terms of value distribution, this means that under market demand $D^{\eta}_{\pi,k}$, η share of consumers have value $v = 0$ and $\pi/(1-k)$ share of them have value $v = 1$. Furthermore, the consumers with values between $\pi/\eta + k$ and 1 are distributed in a way such that the monopolist is indifferent among charging any prices $p \in (\pi/\eta + k, 1]$ in upgrade market k. With this definition, we can now state the main result.

Theorem 1. *For any surplus level $s \in (0, 1)$ and for any regular technology $\gamma \in \Gamma$. The set of (s, γ)-efficient demands is nonempty. Furthermore, every (s, γ)-efficient demand is affine-unit-elastic.*

A major implication of Theorem 1 is that, even though $\mathcal{D}_s$ is an infinite dimensional set, characterizing the efficient market demands among those with the same surplus level is essentially a finite dimensional problem. In particular, finding the market demand that maximizes consumer surplus (monopoly profit, resp.) among every possible demand with the same surplus level becomes a finite dimensional problem that can be completely solved.

Specifically, notice that for any affine-unit-elastic demand $D^{\eta}_{\pi,k}$, the monopolist's (smallest) optimal price equals to $\pi/\eta + k$ in all the upgrade markets $c \leq k$; and equals 1 in all the

[12] Note that $D^{\eta}_{\pi,k}(v) = D^{\eta}_{\pi,0}(v - k)$ for all $v \in [\pi/\eta + k, 1]$, which is a (truncated) affine transformation of a demand that has unit-elasticity on $[\pi/\eta, 1]$.

[13] Formally, the inverse demand of D is defined as $D^{-1}(q) := \sup v \in [0, 1] | D(v) \geq q$, for all $q \in [0, 1]$.

upgrade markets $c \in (k, 1]$. As a result, consumer surplus under $D^{\eta}_{\pi,k} \in \mathcal{D}_s$ is

$$\begin{aligned}\Sigma(D^{\eta}_{\pi,k}|\gamma) = \int_0^k \left(\int_{\frac{\pi}{\eta}+k}^1 D^{\eta}_{\pi,k}(v)\,\mathrm{d}v \right) \gamma(\mathrm{d}c) = \gamma(k) \left(\int_0^1 D^{\eta}_{\pi,k}(v)\,\mathrm{d}v - \int_0^{\frac{\pi}{\eta}+k} D^{\eta}_{\pi,k}(v)\,\mathrm{d}v \right) \\ = \gamma(k) \left[s - \eta \left(\frac{\pi}{\eta} + k \right) \right] \\ = \gamma(k)(s - \pi - \eta k)),\end{aligned} \tag{2}$$

whereas the monopolist's profit is[14]

$$\begin{aligned}\Pi(D^{\eta}_{\pi,k}|\gamma) &= \int_0^k \left(\frac{\pi}{\eta} + k - c \right) D^{\eta}_{\pi,k} \left(\frac{\pi}{\eta} + k \right) \gamma(\mathrm{d}c) + \int_k^{\bar{c}} (1-c) D^{\eta}_{\pi,k}(1) \gamma(\mathrm{d}c), \\ &= \int_0^k \left(\frac{\pi}{\eta} + k - c \right) \eta\gamma(\mathrm{d}c) + \int_k^{\bar{c}} (1-c) \frac{\pi}{1-k} \gamma(\mathrm{d}c) \\ &= \int_0^k [\pi + (k-c)\eta] \gamma(\mathrm{d}c) + \int_k^{\bar{c}} (1-c) \frac{\pi}{1-k} \gamma(\mathrm{d}c).\end{aligned} \tag{3}$$

Together with Theorem 1, this means that the efficient frontier given any $s \in (0, 1)$ and any regular $\gamma \in \Gamma$ can be characterized by only three parameters π, k, η with two constraints: $\pi/\eta + k \leq 1$ and $D^{\eta}_{\pi,k} \in \mathcal{D}_s$.

The proof of Theorem 1 can be found in the appendix. In what follows, we outline the main ideas of the proof of Theorem 1. Notice that for any $s \in (0, 1)$ and for any regular $\gamma \in \Gamma$, the set of efficient market demands equals to the solutions of the following class of optimization problems:

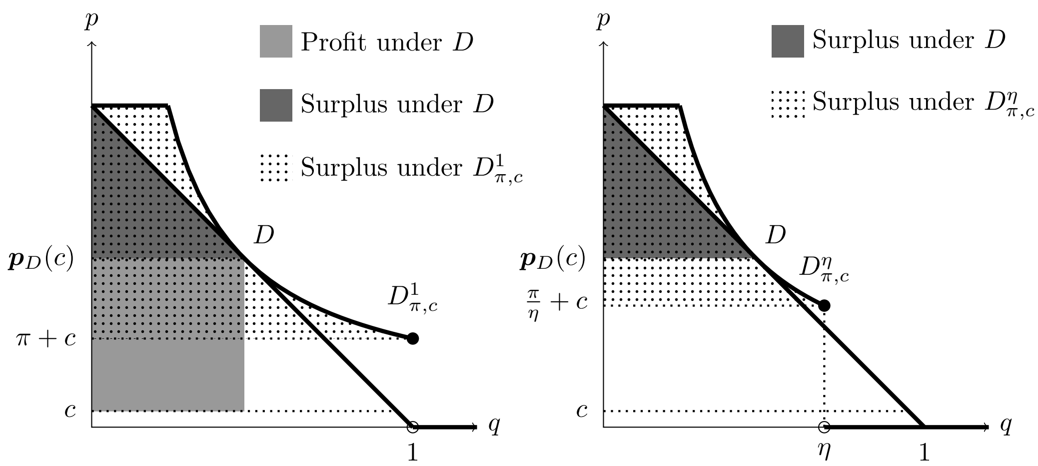
$$\begin{aligned}&\sup_{D \in \mathcal{D}_s} [\alpha \Pi(D|\gamma) + \beta \Sigma(D|\gamma)] \\ =& \sup_{D \in \mathcal{D}_s} \left[\int_0^{\bar{c}} \left(\alpha \cdot \max_{p \geq 0} (p-c) D(p) + \beta \cdot \int_{\boldsymbol{p}_D(c)}^1 D(v)\,\mathrm{d}v \right) \gamma(\mathrm{d}c) \right]\end{aligned} \tag{4}$$

for some $\alpha, \beta \geq 0$ with $\alpha + \beta > 0$. Therefore, characterizing the set of efficient market demands is equivalent to solving (4) for every $\alpha, \beta \geq 0$ with $\alpha + \beta > 0$.

To better understand (4), we first consider a relaxed problem, where the market demands are allowed to be conditioned on each upgrade market. That is, instead of finding a market demand $D \in \mathcal{D}_s$ that maximizes the mixture of the weighted sum of monopoly profit and consumer surplus, we first solve for the market demand that maximizes the weighted sum in a given upgrade market $c \in [0, 1]$. This relaxed problem is studied by Roesler and Szentes (2017). In fact, they show that for any $s \in (0, 1)$ and for any upgrade market $c \in [0, 1]$, an affine-unit-elastic demand $D^{\eta}_{\pi,c}$ solves this relaxed problem.

[14] As a notational convention, we write $\int_a^b f(x)\,\mathrm{d}x = 0$ for any integrable function f whenever $a > b$.

Figure 2: Upgrade Market c



(a) D and $D^1_{\pi,c}$ (b) $D^\eta_{\pi,c} \in \mathcal{D}_s$

To see this, consider any market demand $D \in \mathcal{D}_s$ and any upgrade market $c \in [0, 1]$.[15] The monopoly profit and consumer surplus under this demand in upgrade market c are depicted by the shaded areas in Figure 2a. Let $\pi := (\boldsymbol{p}_D(c) - c)D(\boldsymbol{p}_D(c))$ be the monopoly profit under D in upgrade market c. Optimality of $\boldsymbol{p}_D(c)$ implies that

$$(v - c)D(v) \leq (\boldsymbol{p}_D(c) - c)D(\boldsymbol{p}_D(c)) = \pi \iff D(v) \leq \frac{\pi}{v - c},$$

for all $v \in (c, 1]$. Therefore, as depicted in Figure 2a, the affine-unit-elastic demand $D^1_{\pi,c}$ is always above D. Moreover, any price in the interval $[\pi + c, 1]$ generates the same profit for the monopolist in upgrade market c under the affine-unit-elastic demand $D^1_{\pi,c}$. In other words, we may regard this affine-unit-elastic demand as an *iso-profit curve* for the monopolist in upgrade market c, where the parameter π describes the profit level. From this perspective, the optimal price $\boldsymbol{p}_D(c)$ in market c is given by the price at which market demand D is tangent to the iso-profit curve $D^1_{\pi,c}$ and the optimal profit level is π.

Of course, the affine-unit-elastic demand $D^1_{\pi,c}$ does not have surplus level s. Nonetheless, by the intermediate value theorem, there exists $\eta \in [0, \pi/(1 - c)]$ such that the affine-unit-elastic demand $D^\eta_{\pi,c}$ has surplus s, as depicted in Figure 2b.[16] Furthermore, under the affine-unit-elastic demand $D^\eta_{\pi,c}$, every price in the interval $[\pi/\eta + c, 1]$ is optimal for the monopolist

[15] Upgrade markets $c > 1$ are trivial since the highest possible value is always 1 for any market demand $D \in \mathcal{D}_s$ and thus all of these market have no trade.

[16] To see this, notice that $D^1_{\pi,c}$ has surplus higher than s since it is pointwise above $D \in \mathcal{D}_s$. Meanwhile, notice that since the function $c \mapsto \max_p (p - c)D(p)$ is convex and since $\max_p pD(p) \leq s$ while $\max_p (p - 1)D(p) = 0$, it must be that $\max_p (p - c)D(p) \leq (1 - c)s$ for all $c \in [0, 1]$. Thus, $\pi/(1 - c) \leq s$, which in turn implies that $G^{\frac{\pi}{1-c}}_{\pi,c}$ has surplus level $\pi/(1 - c) \leq s$. Therefore, by the intermediate value theorem, since the surplus level induced by $G^{\eta'}_{\pi,c}$ is continuous in η', there exists η such that $G^\eta_{\pi,c}$ induces surplus level s.

in upgrade market c and yields profit π. Meanwhile, as shown by Roesler and Szentes (2017) and demonstrated in Figure 2b, consumer surplus becomes higher under $D^{\eta}_{\pi,c}$.[17] As a result, changing the demand from D to $D^{\eta}_{\pi,c}$ increases consumer surplus while leaving the monopolist's profit unchanged in upgrade market c. As a result, in a fixed upgrade market c, any efficient market demand must be payoff-equivalent to an affine unit elastic demand $D^{\eta}_{\pi,c} \in \mathcal{D}_s$, for some π, η.

Essentially, the argument above means that the monopolist's iso-profit curves (and hence the affine-unit-elastic demands) contain sufficient information to trace out the entire efficient payoff frontier *when conditioning on a particular upgrade market*. Nevertheless, this argument does not hold once we return to (4) from the relaxed problem. Specifically, since both the monopoly profit and consumer surplus are mixtures of those in a continuum of upgrade markets, and since the market demands are not allowed to be conditioned on individual upgrade markets, the collection of affine-unit-elastic demands cannot be regarded as iso-profit curves of the monopolist. After all, the monopolist has to consider prices in *all* upgrade markets at the same time.[18] Therefore, Theorem 1 does *not* immediately follow from characterizations of efficient market demands in each upgrade market. In fact, part of the significance of Theorem 1 is exactly that affine-unit-elastic demands are still sufficient for tracing out the efficient payoff frontier, even though they do not serve as the monopolist's iso-profit curves. The essence of the proof relies on a local perturbation argument, which we will now explain.

The proof of Theorem 1 is based on the following claim: Given any $s \in (0,1)$ and any regular $\gamma \in \Gamma$. If a market demand $D \in \mathcal{D}_s$ is not affine-unit-elastic, then we can always perturb D and construct another market demand $\widehat{D} \in \mathcal{D}_s$ such that $\Pi(\widehat{D}|\gamma) \approx \Pi(D|\gamma)$ and $\Sigma(\widehat{D}|\gamma) > \Sigma(D|\gamma)$. Together with the property that $\Pi(\cdot|\gamma)$ and $\Sigma(\cdot|\gamma)$ are continuous and upper-semicontinuous, respectively (see Corollary 1 of Yang (2020)), it then follows that any efficient market demand must be affine-unit-elastic.

Here we describe the perturbation. Consider any market demand $D \in \mathcal{D}_s$. To begin with, we first note that it is without loss to assume that D is constant on $(0, \boldsymbol{p}_D(0)]$. Indeed, if not, then as Figure 3 shows, there exists $\hat{\eta} \in [D(\boldsymbol{p}_D(0)), 1]$ such that for $\widehat{D}$ defined as

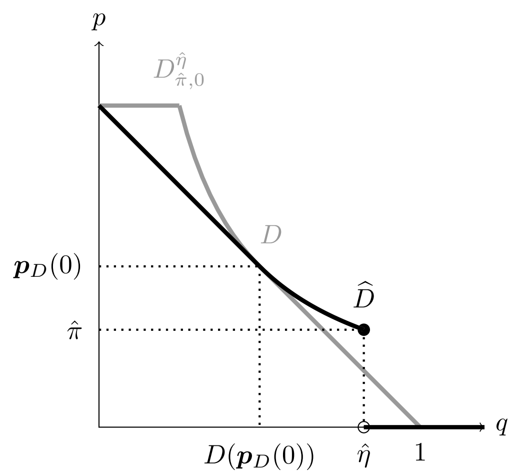
$$\widehat{D}(v) := \begin{cases} D^{\hat{\eta}}_{\hat{\pi},0}(v), & \text{if } v \in [0, \boldsymbol{p}_D(0)] \\ D(v), & \text{if } v \in (\boldsymbol{p}_D(0), \infty) \end{cases},$$

$\widehat{D} \in \mathcal{D}_s$, where $\hat{\pi} := \boldsymbol{p}_D(0) D(\boldsymbol{p}_D(0))$. Under the market demand $\widehat{D}$, the monopolist's optimal

[17] Although Figure 2b only demonstrates the case where $\eta \geq D(\boldsymbol{p}_D(c))$, it can also be shown that consumer surplus increases even when $\eta < D(\boldsymbol{p}_D(c))$, by using the fact that the monopolist's profit remains as π and that $D^{\eta}_{\pi,c} \in \mathcal{D}_s$.

[18] In particular, under any affine-unit-elastic demand $D^{\eta}_{\pi,k}$, the monopolist's (unique) optimal price equals $\pi/\eta + k$ in all upgrade markets $c < k$; and equals to 1 in all upgrade markets $c > k$.

Figure 3: $\widehat{D} \in \mathcal{D}_s$ is Constant on $(0, \boldsymbol{p}_D(0)]$



price in each upgrade market $c > 0$ remains the same, and hence the monopoly profit and consumer surplus remain the same in each upgrade market $c > 0$. By regularity of γ (in particular, that $\gamma(0) = 0$), it then follows that $\Pi(\widehat{D}|\gamma) = \Pi(D|\gamma)$ and that $\Sigma(\widehat{D}|\gamma) = \Sigma(D|\gamma)$. That is, D and $\widehat{D}$ are payoff-equivalent. Henceforth, we assume that D is constant on $[0, \boldsymbol{p}_D(0)]$.

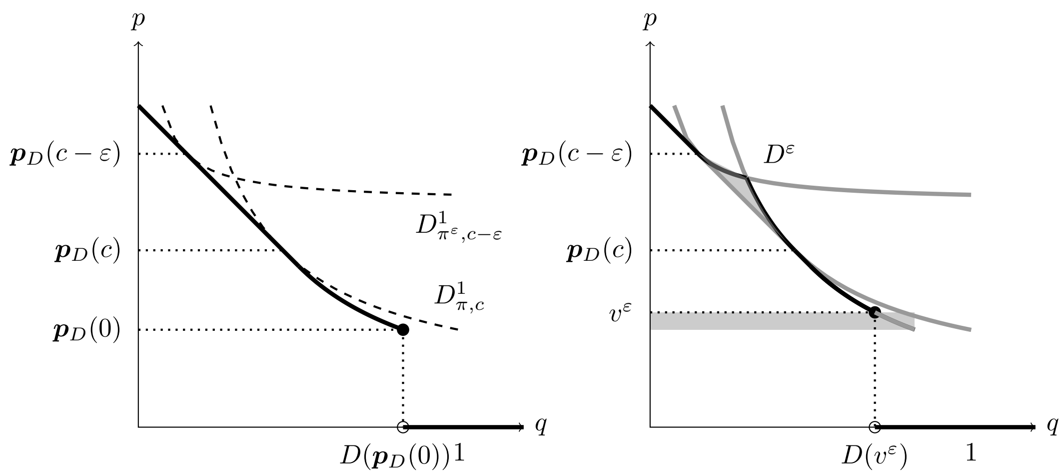
Since D is not affine-unit-elastic, there must be upgrade markets $0 < c_1 < c_2$ in which the monopolist has distinct optimal prices: $\boldsymbol{p}_D(0) < \boldsymbol{p}_D(c_1) < \boldsymbol{p}_D(c_2)$. To highlight the main insights and avoid unnecessary complications, we further assume that under this market demand D, the monopolist's optimal price is continuous and strictly increasing in upgrade markets c.[19] The perturbation of other types of market demands can be found in the appendix. For any such market demand $D \in \mathcal{D}_s$, consider any two nearby upgrade markets $0 < c - \varepsilon < c$. The iso-profit curves for the monopolist under her optimal prices in markets c and $c - \varepsilon$ are depicted in Figure 4a (labeled as $D^1_{\pi,c}$ and $D^1_{\pi^\varepsilon, c-\varepsilon}$, respectively). Since the iso-profit curves under optimal prices are always above the market demand, it must be that $\min\{D^1_{\pi,c}(v), D^1_{\pi^\varepsilon,c-\varepsilon}(v)\} \geq D(v)$ for all $v \in [\boldsymbol{p}_D(c-\varepsilon), \boldsymbol{p}_D(c)]$. Therefore, for each $\varepsilon > 0$, there exists $v^\varepsilon \in [\boldsymbol{p}_D(0), \boldsymbol{p}_D(c-\varepsilon)]$ such that for D^ε defined as

$$D^\varepsilon(v) := \begin{cases} 1, & \text{if } v = 0 \\ D(v^\varepsilon), & \text{if } v \in (0, v^\varepsilon] \\ \min\{D^1_{\pi,c}(v), D^1_{\hat{\pi},\hat{c}}(v)\}, & \text{if } v \in (\boldsymbol{p}_D(c), \boldsymbol{p}_D(\hat{c})] \\ D(v), & \text{otherwise} \end{cases},$$

$D^\varepsilon \in \mathcal{D}_s$. The perturbed demand D^ε is depicted in Figure 4b.

[19]This is equivalent to saying that the right-limit of $1 - D$, which is a CDF, is regular in the Myersonian sense, which in turn is equivalent to saying that the marginal revenue curve induced by D is strictly decreasing.

Figure 4: Local Perturbation



(a) $D^1_{\pi,c}$ and $D^1_{\pi^\varepsilon,c-\varepsilon}$ (b) Perturbed Demand D^ε

Clearly, $\{D^\varepsilon\} \to D$ as $\varepsilon \to 0$ (under the L^1-norm) since $\boldsymbol{p}_D$ is continuous. Furthermore, for any $\alpha, \beta \geq 0$, let

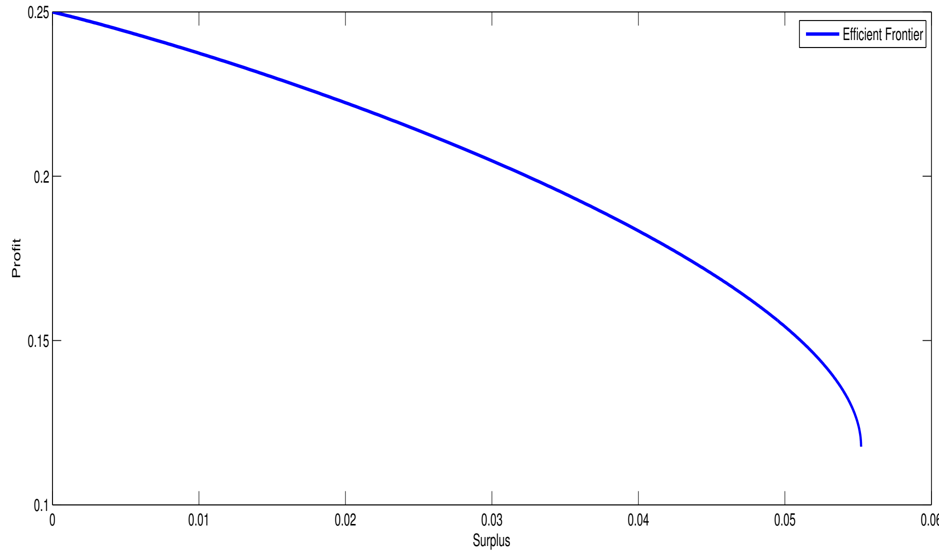
$$\Delta(\varepsilon) := \alpha[\Pi(D^\varepsilon|\gamma) - \Pi(D|\gamma)] + \beta[\Sigma(D^\varepsilon|\gamma) - \Sigma(D|\gamma)]$$

denote the welfare gain when moving from D to D^ε. By construction, $\lim_{\varepsilon\downarrow 0} \Delta(\varepsilon) = \Delta(0) = 0$, and hence it suffices to show that $\Delta'(0) > 0$. In fact, this can be seen from Figure 4b. For any $\varepsilon > 0$, the monopolist's optimal price equals to v^ε in all upgrade markets $c' \in [0, \boldsymbol{p}_D^{-1}(v^\varepsilon)]$; equals to a price between $[\boldsymbol{p}_D(c-\varepsilon), \boldsymbol{p}_D(c)]$ in upgrade markets $c' \in [c-\varepsilon, c]$; and equals to $\boldsymbol{p}_D(c')$ in all other upgrade markets c'. Therefore, the monopolist's profit remain unchanged in every upgrade market $c' \notin [c-\varepsilon, c] \cup [0, \boldsymbol{p}_D^{-1}(v^\varepsilon)]$. Meanwhile, the consumer surplus decreases by an amount that is at most the area of the shaded rectangle in Figure 4b in upgrade markets $c \in [0, \boldsymbol{p}_D^{-1}(v^\varepsilon)]$; and increases by the area of the shaded region below $\min\{D^1_{\pi,c}, D^1_{\pi^\varepsilon,c-\varepsilon}\}$ and above D in every upgrade markets $c' \in (\boldsymbol{p}_D^{-1}(v^\varepsilon), c-\varepsilon]$. Therefore, as $\varepsilon \to 0$ (and hence $\boldsymbol{p}_D^{-1}(v^\varepsilon) \to 0$), the change in monopolist's profit, as well as the loss in consumer surplus vanish in a faster order than that of the gain in consumer surplus, since the upgrade markets in which consumer surplus increases has positive weight under γ (i.e. $\gamma(c) > 0$) in the limit, whereas the weight of upgrade markets in which consumer surplus decreases and monopolist's profit changes converges to zero. This implies that $\Delta'(0) > 0$.

We conclude this section by a numerical example. Suppose that $s = 1/2$, $\gamma(c) = c$ for all $0 \leq c \leq \bar{c} = 1$. By Theorem 1, characterizing efficient market demands is equivalent to finding affine-unit-elastic demands $D^\eta_{\pi,k}$ with surplus level $1/2$ to maximize

$$\alpha\Pi(D^\eta_{\pi,k}|\gamma) + \beta\Sigma(D^\eta_{\pi,k}|\gamma)$$

Figure 5: Efficient Payoff Frontier



for all $\alpha, \beta \geq 0$ with $\alpha + \beta > 0$. According to (2) and (3), this is equivalent to

$$\max_{\pi,k,\eta} \quad \alpha \left[\int_0^k [\pi + (k-c)\eta] \, \mathrm{d}c + \int_k^1 \frac{\pi}{1-k}(1-c) \, \mathrm{d}c \right] + \beta \left[k \left(\frac{1}{2} - \pi - \eta k \right) \right]$$

$$\text{s.t.} \quad \int_0^1 D^{\eta}_{\pi,k}(v) \, \mathrm{d}v = \pi + \eta k + \pi \log \left(\frac{(1-k)\eta}{\pi} \right) = \frac{1}{2}, \qquad \frac{\pi}{\eta} + k \leq 1.$$

Table 1 below summarizes the solution for the case of $\alpha = 0$ and compares it with the solutions conditional on upgrade markets $c = 0$ and $c = 1/2$ (see the online appendix of Roesler and Szentes (2017)). Figure 5 further plots the entire efficient payoff frontier.

Table 1: Comparing Consumer-Optimal Demands

	$c = 0$	$c = \frac{1}{2}$	Multi-Product
π	≈ 0.19	≈ 0.08	≈ 0.09
k	0	0.5	≈ 0.44
η	1	≈ 0.61	≈ 0.63
Surplus	≈ 0.31	≈ 0.12	≈ 0.06

5 Implications on Welfare and Allocation

In this section, we discuss some further implications of Theorem 1. As mentioned above, the significance of Theorem 1 is that characterizing the efficient frontier can be reduced to to a finite dimensional problem, which in turn allows for further comparative statics and welfare analyses, as well as characterizations of market outcomes.

5.1 (Almost) Inevitable Deadweight Loss

One of the most salient features of a monopoly market is that it is inefficient. As the monopolist would always charge a mark-up, deadweight loss is an ubiquitous concept in the theory of monopoly. However, a surprising result of Roesler and Szentes (2017) is that when the monopolist sells a single product at a constant marginal cost, the deadweight loss is always zero under any efficient demand (among those with the same surplus level, see Section 3 of their Online Appendix). The characterization of Theorem 1 allows us to explore the same question in the context of multi-product monopoly. As shown in Proposition 1 below, the same conclusion does not hold in a multi-product monopoly. There is (almost) always a positive amount deadweight loss in a multi-product monopoly even under the efficient market demands. To formalize the claim, we define, for any market demand D and for any regular technology $\gamma \in \Gamma$, the amount of deadweight loss as follows:

$$L(D,\gamma) := \int_0^{\bar{c}} \left[\int_c^1 D(v)\,\mathrm{d}v - (\boldsymbol{p}_D(c) - c)D(\boldsymbol{p}_D(c)) - \int_{\boldsymbol{p}_D(c)}^1 D(v)\,\mathrm{d}v \right] \gamma(\mathrm{d}c).$$

That is, the amount of deadweight loss under market demand D and technology γ is the mixture of deadweight losses in each upgrade market c. Clearly, $L(D,\gamma) \geq 0$. With this notation, we can formally state the following result:

Proposition 1. *Consider any surplus level $s \in (0,1)$ and any regular technology $\gamma \in \Gamma$. For any (s,γ)-efficient market demand D, $L(D,\gamma) = 0$ if and only if $D(v) = s$ for all $v \in (0,1]$.*

Proof. By Theorem 1, any (s,γ)-efficient demands D must be affine-unit-elastic. Therefore,

$$\begin{aligned} L(D,\gamma) &= \int_0^{\bar{c}} \left[\int_c^1 D^\eta_{\pi,k}(v)\,\mathrm{d}v - (\boldsymbol{p}_{D^\eta_{\pi,k}}(c) - c)D(\boldsymbol{p}_{D^\eta_{\pi,k}}(c)) - \int_{\boldsymbol{p}_{D^\eta_{\pi,k}}(c)}^1 D^\eta_{\pi,k}(v)\,\mathrm{d}v \right] \gamma(\mathrm{d}c) \\ &= \int_0^k [s - c\eta - \pi - (s - \pi - c\eta)]\,\gamma(\mathrm{d}c) + \int_k^{\frac{\pi}{\eta}+k} [s - c\eta - \frac{\pi}{1-k}(1-c)]\gamma(\mathrm{d}c) \\ &\quad + \int_{\frac{\pi}{\eta}+k}^{\bar{c}} \left[\int_{\frac{\pi}{\eta}+k} \frac{\pi}{v-k}\,\mathrm{d}v - \frac{\pi}{1-k}(1-c) \right] \gamma(\mathrm{d}c) \\ &\geq 0, \end{aligned}$$

for some π, η, k such that $D^\eta_{\pi,k} \in \mathcal{D}_s$ and that $\pi/\eta + k \leq 1$, with equality if and only if $\pi = 0$, $\eta = s$ and $k = 1$. This completes the proof. ■

The essence of Proposition 1 is closely related to the comparison between Roesler and Szentes (2017) and this paper. As discussed in Section 4, when market demands are allowed to be conditioned on each upgrade market, the affine-unit-elastic demands coincide with the iso-profit curves of the monopolist in that market. As a result, one can always hold the monopoly profit and total surplus level fixed while reducing the amount of deadweight loss if the market demand is not affine-unit-elastic. This argument is not valid when market demands cannot be conditioned on upgrade markets. As a result, even if efficient market demands are still affine-unit-elastic, there would still be deadwieght losses in almost all upgrade markets since a given affine-unit-elastic demand only corresponds to an iso-profit curve of the monopolist in one upgrade market. The only exception is when the market demand is constant on $(0, 1)$ (i.e., when the consumers' values are concentrated at 0 or 1), in this case the monopolist's optimal price equals to 1 in every upgrade market and the consumer surplus is completely extracted.

Consequently, Proposition 1 can be viewed as "bringing back" deadweight losses to monopolistic markets. According to Proposition 1, in a multi-product monopoly, deadweight losses are inevitable even under the efficient market demands, in contrast to a single product monopoly with constant marginal cost as in Roesler and Szentes (2017).

5.2 The Effect of Production Technology

As motivated in the introduction, it is crucial to understand the effect of production technologies on market outcomes, in addition to the effects of market demands. Just as market demands, production technologies are also infinite dimensional. Nevertheless, since Theorem 1 reduces the efficient market demands to a finite-dimensional family, it makes comparative statics of production technologies tractable. We first consider the effects of changes in technology on consumer surplus.

As an immediate corollary of Theorem 1 and (2), characterizing the market demand $D \in \mathcal{D}_s$ that maximizes consumer surplus is equivalent to the following (finite dimensional) constraint optimization problem:

$$\begin{aligned}&\max_{\pi,k,\eta} \gamma(k)(s - \pi - \eta k)\\ &\text{s.t. } \pi + \eta k + \pi \log\left(\frac{(1-k)\eta}{\pi}\right) = s, \quad \frac{\pi}{\eta} + k \leq 1,\end{aligned}$$

which in turn, by letting

$$\begin{aligned}\omega(k) :=& \min_{\pi,\eta}[\pi + \eta k]\\ &\text{s.t. } \pi + \eta k + \pi \log\left(\frac{(1-k)\eta}{\pi}\right) = s, \quad \frac{\pi}{\eta} + k \leq 1,\end{aligned}$$

for all $k \in [0,1]$, can be written as

$$\max_{k\in[0,1]} \gamma(k)(s-\omega(k)). \tag{5}$$

Furthermore, it is straightforward to show that ω is continuous on $[0,1]$, differentiable on $(0,1)$ and strictly increasing on $[0,1]$ with $\omega(1)=s$. As a result, an equivalent way to write (5) is

$$\max_{p\in[\omega(0),s]} \gamma\circ\omega^{-1}(p)(s-p). \tag{6}$$

Thus, (6) suggests that maximizing buyer's surplus is in fact equivalent to maximizing a monopsonist's surplus whose value of the good is s and is facing a supply function $\gamma\circ\omega^{-1}$. This observation, together with the insight of Johnson and Myatt (2006b), yields an unambiguous comparative statics result. More specifically, consider a family of regular technologies $\{\gamma_\lambda|\lambda \in [0,1]\} \subseteq \Gamma$ such that $\partial\gamma_\lambda(c)/\lambda$ exists for all $\lambda\in[0,1]$ and for all $c \geq 0$. Define an ordering on $\{\gamma_\lambda|\lambda\in[0,1]\}$ as the following.

Definition 2. A family $\{\gamma_\lambda|\lambda\in[0,1]\}\subseteq\Gamma$ is said to be ranked by the *rotational order* if there exists $\{c_\lambda|\lambda\in[0,1]\}\subseteq[0,1]$ such that c_λ is nondecreasing in λ and that

$$\frac{\partial\gamma_\lambda(c)}{\partial\lambda} \leq 0 \text{ if } c < c_\lambda, \quad \text{and} \quad \frac{\partial\gamma_\lambda(c)}{\partial\lambda} \geq 0 \text{ if } c > c_\lambda, \; \forall\lambda\in[0,1].$$

Motivated by the analyses of Johnson and Myatt (2006b), as a corollary of Theorem 1, the following comparative statics can then be derived.

Proposition 2. *For any family of regular technologies $\{\gamma_\lambda|\lambda\in[0,1]\}\subseteq\Gamma$ ranked by the rotational order, the consumer-optimal surplus is quasi-convex in λ.*

Proof. By Theorem 1 and (5), for any $\lambda\in[0,1]$,

$$\sup_{D\in\mathcal{D}_s}\Sigma(D|\gamma_\lambda) = \max_{k\in[0,1]}\gamma_\lambda(k)(s-\omega(k)),$$

Furthermore, since $\{\gamma_\lambda|\lambda\in[0,1]\}$ is ranked by the rotational order, for any $k\geq 0$, the function $\lambda\mapsto\gamma_\lambda(k)(s-\omega(k))$ is quasi-convex. Therefore, as a function of λ, the consumer-optimal surplus is quasi-convex in λ since it is a pointwise supremum of a family of quasi-convex functions. This completes the proof. ∎

Proposition 2 means that for any family of technologies $\{\gamma_\lambda|\lambda\in[0,1]\}\subseteq\Gamma$ ranked by the rotational order, the consumer-optimal surplus as a function of the parameter λ is either increasing, decreasing, or U-shaped. A crucial implication is that consumer-optimal surplus is nonmonotonic in cost function in general. Specifically, recall that a regular technology $\gamma\in\Gamma$ is uniquely associated with a strictly convex cost function C with $C(0)=0$ through

(1). Thus, if γ_0 is associated with C_0 and γ_1 is associated with C_1, then γ_0 second order stochastically dominates γ_1 if and only if $C_1 \leq C_0$ pointwise.[20] Therefore, for a family of technologies $\{\gamma_\lambda | \lambda \in [0,1]\}$ ranked by second order stochastic dominance that are also consistent with the rotational order, it is possible that consumer surplus would be higher under a larger cost function (lower λ) then under a smaller cost function (higher λ). In fact, quasi-convexity implies that consumer surplus must the highest under either the largest cost function or the smallest cost function.

In fact, the nonmonotonicity property is not specific to only consumer-optimal surplus. As Proposition 3 below shows, (almost) all the optimal weighted sum of monopoly profit and consumer surplus could be nonmonotonic in cost function. To formalize this claim, let

$$W^{\alpha,\beta}(s,\gamma) := \sup_{D \in \mathcal{D}_s} \quad [\alpha \Pi(D|\gamma) + \beta \Sigma(D|\gamma)],$$

for any surplus level $s \in (0,1)$, any technology $\gamma \in \Gamma$ and for any weights $\alpha, \beta \geq 0$ with $\alpha + \beta > 0$.

Proposition 3. *For any surplus level $s \in (0,1)$, for any regular technology $\gamma \in \Gamma$, and for any weights $\alpha \geq 0$, $\beta > 0$, there exists a regular technology $\tilde{\gamma}$ which second order stochastically dominates γ such that $W^{\alpha,\beta}(s,\gamma) > W^{\alpha,\beta}(s,\tilde{\gamma})$.*

Proof. See appendix. ∎

Proposition 3 essentially means that, for any surplus level $s \in (0,1)$ and any regular technology $\gamma \in \Gamma$, there exists another (nearby) technology under which the efficient frontier shifts "inward", even though the cost of producing every quality level decreases.[21] This suggests that reduction of cost may not always enhance efficiency. In fact, there might even be cases where reduction of cost leads to efficiency loss.

5.3 Market Outcomes under Efficient Demands

In addition to the monopoly profit and consumer surplus, we may also examine the allocative outcomes of the market. That is, the distribution of products across consumers. In general,

[20]By Theorem 3.8 of Sriboonchita et al. (2009),

$$\int_0^k \gamma_1(c)\,\mathrm{d}c \geq \int_0^k \gamma_0(c)\,\mathrm{d}c, \forall k \geq 0 \iff C_0(q) = \int_0^q C_0'(x)\,\mathrm{d}x \leq \int_0^q C_1'(x)\,\mathrm{d}x = C_1(q), \forall q \in [0,1].$$

[21]It is noteworthy that Proposition 3 does *not* imply that for any payoff pair on the efficient frontier, there exists another technology such that an outcome under this new technology Pareto-dominates the original one. Instead, it only implies there exits an other technology that exhibits higher costs and yet the original payoff pair falls below the frontier locally.

allocative outcomes in a multi-product monopoly with quality-differentiated products could be quite complex, especially when the monopolist can produce a continuum of different quality levels. After all, as shown by Mussa and Rosen (1978), this is often a nontrival nonlinear pricing problem. Alternatively, using the language of Johnson and Myatt (2003), Johnson and Myatt (2006a) and Johnson and Myatt (2006b) (and hence that of this paper), the monopolist's optimal price (and hence the share of purchasing consumers) could be different in every upgrade market c. While the continuum-quality model has the benefit of being tractable, the complexity of its implied market outcomes are sometimes less desirable from both a theoretical and an empirical point of view.

However, it turns out that under any efficient market demand, the allocative outcome is actually quite simple, as shown by Proposition 4 below.

Proposition 4. *For any surplus level $s \in (0,1)$ and for any regular technology $\gamma \in \Gamma$, under any (s,γ)-efficient market demand, only two quality levels are sold. That is, there exists $k \in [0,1]$ such that the monopolist optimally bundles upgrades $c \in [0,k]$ and $c \in (k,1]$, and then sells these bundles at two prices.*

Proof. By Theorem 1, any (s,γ)-efficient market demand must equal to some affine-unit-elastic demand $D^{\eta}_{\pi,k} \in \mathcal{D}_s$. Moreover, under $D^{\eta}_{\pi,k}$, the monopolist's (lowest) optimal price equals to $\pi/\eta + k$ in all upgrade markets $c \in [0,k]$, and equals to 1 in all upgrade markets $c \in (k,1]$. This completes the proof. ■

According to Proposition 4, although allocative outcomes in a multi-product monopoly are generally complex, they are rather simple on the efficient frontier. Under any efficient market demand, the monopolist only sells two bundles of upgrades, one "basic" level (i.e., $c \in [0,k]$) and the other "advanced level" (i.e., $c \in (k,1]$). Even though the technology allows the monopolist to produce a continuum of quality-differentiated products, the monopolist would end up selling only two different products under any efficient market demand. Another way to view this result is that complexity of the allocative outcome and efficiency are unrelated. Observing simple allocative outcome does not necessarily mean that the market demand is less efficient. Conversely, a market with a complex allocative outcome is in fact inefficient.

6 Discussions

6.1 Interpretations of the Model

Throughout the paper, we interpret the model as a nonlinear pricing problem where a multi-product monopoly sells quality-differentiated products to a unit mass of consumers, and characterize the efficient market demands among those that have the same surplus level. From the perspective of Johnson and Myatt (2003), Johnson and Myatt (2006a), and Johnson and Myatt (2006b) we transform this problem into a continuum of single-product monopoly pricing problems with constant marginal costs, where each marginal cost level represents an upgrade market. Mathematically, this "upgrade approach" is in turn equivalent to a model where the monopolist's marginal cost is uncertain. Therefore, Theorem 1 above can be regarded as a characterization of the ex-ante efficient demands among those with the same surplus level; Proposition 2 and Proposition 3 can be regarded as nonmonotonicity of consumer surplus and welfare when the monopolist's cost becomes more uncertain (in the sense of mean-preserving spread); and Proposition 4 can be viewed from the perspective of ex-post allocation under efficient market demands.

Meanwhile, although we interpret the demand functions as the distribution over consumers values and search for the efficient demands among those with the same surplus level, as hinted by the comparison with Roesler and Szentes (2017) above, the model is equivalent to the setting where there is one buyer with binary value $v \in \{0, 1\}$ that is distributed according to $(1 - s, s)$ and can receive different signals (i.e., Blackwell experiments) for this value. Under this interpretation, a demand function $D \in \mathcal{D}_s$ can be regarded as a posterior distribution of value that is a mean-preserving contraction of the prior $(1-s, s)$, and hence a Blackwell experiment. Therefore, Theorem 1 can also be interpreted as the characterization of efficient signals. From this perspective, Theorem 1 also leads to a characterization of bidders' best response in an auction game where bidders choose signals simultaneously before the seller designs an optimal auction. In an earlier version of this paper, we completely characterize the unique symmetric equilibrium of the game using Theorem 1, where each bidder randomizes over signals that induce posteriors of form $D^{\eta}_{\pi,k}$.

6.2 More Restrictive Set of Demands

When characterizing the efficient market demands, we search across all market demands with the same surplus level. While allowing for all such demand functions render an infinite dimensional problem, the richness of this set also brings tractability. Specifically, when constructing a local perturbation of any non-affine-unit-elastic demands, the only constraint the perturbed demand has to satisfy is that it must have the same surplus level as the original one. This is a crucial reason for why we are able to obtain a complete characterization. Of course, a natural extension would be to impose more restrictions on the set of market demands that are allowed. One example would be to impose a mean-preserving constraint so that any feasible demand must be a mean-preserving contraction of a given demand D_0.[22] For instance, there could be some exogenous limits on how spread out the distribution of consumers values could be, such as underlying preferences, income distribution, or a common prior that has non-binary support.

[22] When $D_0(v) = s$ for all $v \in (0, 1]$, this restriction reduces to the surplus level constraint above.

With an arbitrary upper bound under the mean preserving spread order, characterizing efficient market demands becomes much more difficult. After all, the maximization problem (4) has only one constraint when searching across all demands with the same surplus level. In contrast, there would be a continuum of constraints when searching across all demands that are mean preserving contractions of some arbitrary demand. In particular, for arbitrary D_0, the local perturbation constructed in Section 4 would not be a mean preserving contraction of D_0 in general, even if it has the same surplus level as D_0. Nonetheless, using similar local perturbation arguments, the consumer-optimal demand can be shown to exhibit the following feature: There exists an increasing sequence $\{v_n\}$ that forms a partition of $[0, 1]$, such that in the interior of each element of this partition, (v_n, v_{n+1}), the demand is affine-unit-elastic, while the mean preserving spread constraint binds at v_n and v_{n+1}.[23] Nonetheless, the class of demands of this form could still be infinite dimensional (as the sequence $\{v_n\}$ might be countably infinite), and hence complete characterization remains an open question.

6.3 Policy Implications

The results about the effect of market demands and production technologies on market outcomes further lead to some policy implications. From the demand side, Theorem 1, Proposition 1, and Proposition 4 provide some insight about demand manipulation. As motivated in Section 1, and as demonstrated by Johnson and Myatt (2006b), plenty of economic activities can affect the shape of market demands (e.g.,product design, information disclosure, income redistribution, or marketing strategies). Therefore, for a "demand designer" (e.g., monopolist, regulator, or consumers) who can affect market demands through these activities, Theorem 1 prescribes the exact shape for a market demand to be efficient. Furthermore, from a policymaker's point of view, Proposition 4 implies that if the observed allocative outcome is more complicated than a simple two-products market, then the market can certainly be Pareto-improved via (proper) demand manipulations. Nevertheless, Proposition 1 indicates that even the efficient market demands are not able to eliminate deadweight loss, and hence other types of intervention must be needed if a policymaker's goal is to eliminate deadweight loss.

On the technology side, Proposition 2 and Proposition 3 indicate that in terms of efficiency (and consumer surplus in particular), the *level* of the monopolist's cost alone is not indicative. Instead, it is the *curvature* of the technology γ that has crucial implications. Indeed, both the optimal consumer surplus and the optimal weighed sum of monopoly profit and consumer

[23] This is reminiscent of the *monotone partition* structure in the literature, see, for instance Kolotilin, Li, Mylovanov, and Zapechelnyuk (2017), Dworczak and Martini (2019), and Kleiner, Moldovanu, and Strack (forthcoming)), except that when the mean preserving constraint does not bind, the demand is affine-unit-elastic as opposed to being constant almost everywhere)

surplus are nonmonotonic in the level of cost function. As a result, from a policymaker's perspective, advancement of technology in the sense of reduction of cost may not necessarily be desirable in a multi-product monopoly. This observation could provide insights for patent regulation and vertical integration.

7 Conclusion

In this paper, we characterize the efficient market demands with a fixed surplus level in a multi-product monopoly where the monopolist sells a continuum of quality-differentiated products. By the insight of Johnson and Myatt (2003), Johnson and Myatt (2006a) and Johnson and Myatt (2006b), we transform the monopolist's problem into a continuum of single-product pricing problem across different upgrade markets, where the monopolist has different (constant) marginal costs in each upgrade market. We then use a local perturbation argument to show that any efficient market demand must be affine-unit-elastic. This characterization has several implications. First, it reduces the problem of finding the efficient demands to a finite dimensional problem, even though the set of demand functions with the same surplus level is infinite-dimensional. Secondly, it implies that under (almost) every efficient demand, there is still a positive amount of deadweightloss, in contrast to those in a single-product monopoly where deadweight loss is always zero. Furthermore, the characterization of efficient market demands implies that only two different products are sold under any efficient market demand. That is, under any efficient demand, the monopolist sells the "upgrades" in only two bundles. Lastly, using this characterization, we are also able to examine how changes in the production technology affect market outcomes. We show that the optimal weighed sum of consumer surplus and monopoly profit is non-monotonic in production cost.

Although the market demands can be shaped by numerous economic activities as motivated in Section 1, allowing for *all* the demands with the same surplus level certainly abstracts away the fine details of how these demands are formed, as well as the practical limitations when forming different demands. A natural extension is to characterize the set of efficient demands among those that satisfy more complicated constraints. One example is the upper bound in terms of mean preserving spread discussed in Section 6. Other examples include lower bound in terms of mean preserving contraction (e.g., intrinsic noises in consumers' tastes); upper/lower bounds in terms of first-order stochastic dominance (e.g., designing product upgrades that can only increase consumers' values); and restrictions on higher-order moments. These can be topics for future research.

Lastly, the methodology developed in this paper is related to a robust pricing problem in a multi-product monopoly, where the monopolist sells a continuum of quality-differentiated products at increasing marginal costs, but does not have any knowledge about the market demand, except the total surplus level. In a single-product monopoly, this would correspond to the model of Carrasco et al. (2018) with only the first moment constraint. The local perturbation method appears to be useful for solving the monopolist's min-max problem, which can be useful in identifying the max-min solution. This can also be a topic for future research.

References

Bergemann, D., B. Brooks, and S. Morris (2015): "The Limits of Price Discrimination," *American Economic Review*, 105, 921–957.

Bergemann, D. and M. Pesendorfer (2007): "Information Structures in Optimal Auctions," *Journal of Economic Theory*, 137, 580–609.

Bergemann, D. and K. Schlag (2011): "Robust Monopoly Pricing," *Journal of Economic Theory*, 146, 2527–2543.

Brooks, B. (2013): "Surveying and Selling: Belief Extraction in Auctions," Working Paper.

Carrasco, V., V. F. Luz, N. Kos, M. Messner, P. Monterio, and H. Moreira (2018): "Optimal Selling Mechanisms under Moment Consitions," *Journal of Economic Theory*, 177, 245–279.

Condorelli, D. and B. Szentes (2020): "Information Design in the Hold-up Problem," *Journal of Political Economy*, 128, 681–709.

——— (2021): "Surplus Bounds in Cournot Monopoly and Competition," Working Paper.

Dworczak, P. and G. Martini (2019): "The Simple Economics of Optimal Persuasion," *Journal of Political Economy*, 127, 1993–2048.

Haghpanah, N. and R. Siegel (2020): "Pareto Improving Segmentation of Multi-product Markets," Working Paper.

——— (2021): "Consumer Surplus and Multi-Product Segmentation," Working Paper.

Johnson, J. P. and D. P. Myatt (2003): "Multi-product Quality Competition: Fighting Brands and Product Line Pruning," *American Economic Review*, 93, 748–774.

——— (2006a): "Multi-product Cournot Oligopoly," *RAND Journal of Economics*, 37, 583–601.

——— (2006b): "On the Simple Economics of Advertising, Marketing, and Product Design," *American Economic Review*, 96, 756–784.

Kleiner, A., B. Moldovanu, and P. Strack (forthcoming): "Extreme Points and Majorization: Economic Applications," *Econometrica*.

Kolotilin, A., M. Li, T. Mylovanov, and A. Zapechelnyuk (2017): "Persuasion of a Privately Informed Receiver," *Econometrica*, 85, 1949–1964.

Lancaster, K. J. (1975): "Socially Optimal Product Differentiation," *American Economic Review*, 65, 567–585.

Lewis, T. R. and D. E. M. Sappington (1991): "Supplying Information to Facilitate Price Discrimination," *International Economic Review*, 35, 309–327.

Libgober, J. and X. Mu (forthcoming): "Information Robustness in Intertemporal Pricing," *Review of Economic Studies*.

Mussa, M. and S. Rosen (1978): "Monopoly and Product Quality," *Journal of Economic Theory*, 18, 301–307.

Neeman, Z. (2003): "The Effectiveness of English Auctions," *Games and Economic Behavior*, 43, 214–218.

Ravid, D., A.-K. Roseler, and B. Szentes (2019): "Learning Before Trading: On the Inefficiency of Ignoring Free Information," Working Paper.

Renou, L. and K. H. Schlag (2010): "Minimax Regret and Strategic Uncertainty," *Journal of Economic Theory*, 145, 264–286.

Roesler, A.-K. and B. Szentes (2017): "Buyer-Optimal Learning and Monopoly Pricing," *American Economic Review*, 107, 2072–2080.

Sriboonchita, S., W.-K. Wong, S. Dhompongsa, and H. Nguyen (2009): *Stochastic Dominance and Applications to Finance, Risk and Economics*, CRC Press.

Varian, H. R. (1980): "A model of sales," *American Economic Review*, 70, 651–659.

Yang, K. H. (2020): "A Note on Topological Properties of Outcomes in a Monopoly Market," Working Paper.

Appendix

A Proof of Theorem 1

To prove Theorem 1, we first note a close connection between the monopolist's optimal prices in each upgrade market and the shape of demand.

Lemma 1. *A market demand D is affine-unit-elastic if and only if the function $\boldsymbol{p}_D : [0,1] \to \mathbb{R}_+$ is a step function and $D(v) = D(\boldsymbol{p}_D(0))$ for all $v \in (0, \boldsymbol{p}_D(0)]$*

Proof. The "only if" part follows immediately from the definition of affine-unit-elastic demands. For the "if" part, since $\boldsymbol{p}_D(1) = 1$, it must be that $\boldsymbol{p}_D(c) = \boldsymbol{p}_D(0)$ for all $c < c^*$ and that $\boldsymbol{p}_D(c) = 1$ for all $c > c^*$, for some $c^* \in [0,1]$. By Proposition 6 of Yang (2020), $\boldsymbol{p}_D$ is lower-semicontinuous and thus $\boldsymbol{p}_D(c^*) = \underline{v}$. Furthermore, $\max(\operatorname{argmax}_v (v - c^*)D(v)) = 1$ since the optimal price correspondence is upper-hemicontinuous by the same proposition. Therefore, for all $v \in [\boldsymbol{p}_D(0), 1]$,

$$(v - c^*)D(v) = \pi \iff D(v) = \frac{\pi}{v - c^*}.$$

for some $\pi \geq 0$. Together with $D(v) = D(\boldsymbol{p}_D(0))$ for all $v \in (0, \boldsymbol{p}_D(0)]$, we have $D(v) = D_{\pi,c^*}^{D(\boldsymbol{p}_D(0))}$, as desired. ∎

Proof of Theorem 1. First, notice that by Corollary 1 of Yang (2020), for any $\gamma \in \Gamma$, $\Pi(\cdot|\gamma)$ is continuous on $\mathcal{D}_s$ and $\Sigma(\cdot|\gamma)$ is upper-semicontinuous on $\mathcal{D}_s$. Therefore, since $\mathcal{D}_s$ is compact under the weak-* topology, (4) has a solution. Thus, it suffices to show that for any $s \in (0,1)$, for any regular $\gamma \in \Gamma$, for any weights $\alpha \geq 0, \beta \geq 0$ with $(\alpha, \beta) \neq (0,0)$, and for any $D \in \mathcal{D}_s$ that is not affine-unit-elastic, there exists $\widehat{D} \in \mathcal{D}_s$ such that $\alpha\Pi(\widehat{D}|\gamma)\beta\Sigma(\widehat{D}|\gamma) > \alpha\Pi(D|\gamma) + \beta\Sigma(\widehat{D}|\gamma)$.

In the case of $\beta = 0$, notice that for any $s \in (0,1)$, for any regular $\gamma \in \Gamma$, and for any $D \in \mathcal{D}_s$,

$$\begin{aligned}(\boldsymbol{p}_D(c) - c)D(\boldsymbol{p}_D(c)) + \int_{\boldsymbol{p}_D(c)}^1 D(v)\,\mathrm{d}v \leq \int_c^1 D(v)\,\mathrm{d}v \\ \leq s(1-c),\end{aligned}$$

for all $c \in [0,1]$, where the last inequality follows from that fact that any $D \in \mathcal{D}_s$ is a mean preserving contraction of the demand $D_{0,1}^s$ (notice that $D_{0,1}^s(v) = s$ for all $v \in (0,1)$). Moreover, the equality holds if and only if $D(v) = s$ for all $v \in (0,1)$. As a result,

$$\Pi(D|\gamma) < \Pi(D_{0,1}^s|\gamma) = \int_0^{\bar{c}} s(1-c)^+ \gamma(\mathrm{d}c),$$

as desired.

Now consider the case when $\beta > 0$. We show that for any $D \in \mathcal{D}_s$ that is not affine-unit-elastic, there exists a sequence of local perturbation $\{D^\varepsilon\} \subseteq \mathcal{D}_s$ of D such that $\frac{\partial}{\partial\varepsilon}\Pi(D^\varepsilon|\gamma) = 0$ and that $\frac{\partial}{\partial\varepsilon}\Sigma(D^\varepsilon|\gamma) > 0$. Since $\beta > 0$, this would then immediately implies that, there exists $\widehat{D} \in \mathcal{D}_s$ such that $\alpha\Pi(\widehat{D}|\gamma) + \beta\Sigma(\widehat{D}|\gamma) > \alpha\Pi(D/\gamma) + \beta\Sigma(D|\gamma)$.

To this end, consider any $D \in \mathcal{D}_s$ that is not affine-unit-elastic. Notice that since $D \in \mathcal{D}_s$,

$$\int_{\boldsymbol{p}_D(0)}^1 D(v)\,\mathrm{d}v + \boldsymbol{p}_D(0)D(\boldsymbol{p}_D(0)) \leq s.$$

We first claim that it is without loss to assume that either $D(\boldsymbol{p}_D(0)) = 1$ or

$$\int_{\boldsymbol{p}_D(0)}^1 D(v)\,\mathrm{d}v + \boldsymbol{p}_D(0)D(\boldsymbol{p}_D(0)) \leq s$$

Indeed, suppose that $D(\boldsymbol{p}_D(0)) < 1$ and that

$$\int_{\boldsymbol{p}_D(0)}^1 D(v)\,\mathrm{d}v + \boldsymbol{p}_D(0)D(\boldsymbol{p}_D(0)) < s. \tag{7}$$

Let $\pi_0 := \boldsymbol{p}_D(0)D(\boldsymbol{p}_D(0))$. Notice that

$$vD(v) \leq \boldsymbol{p}_D(0)D(\boldsymbol{p}_D(0)) = \pi_0 \iff D(v) \leq \frac{\pi_0}{v} \tag{8}$$

for all $v \in (0,1]$. Since

$$\boldsymbol{p}_D(0)D(\boldsymbol{p}_D(0)) + \int_{\boldsymbol{p}_D(0)}^1 D(v)\,\mathrm{d}v < s,$$

it must be that

$$\int_0^{\boldsymbol{p}_D(0)} D(v)\,\mathrm{d}v + \int_0^{\boldsymbol{p}_D(0)} D(\boldsymbol{p}_D(0))\,\mathrm{d}v = \int_0^{\boldsymbol{p}_D(0)} D(v)\,\mathrm{d}v - \pi_0 > 0. \tag{9}$$

Combining (8) and (9), by the intermediate value theorem, there exists $\hat{v} \in (0, \boldsymbol{p}_D(0))$ such that

$$\int_0^{\hat{v}} \frac{\pi_0}{\hat{v}}\,\mathrm{d}v + \int_{\hat{v}}^{\boldsymbol{p}_D(0)} \frac{\pi_0}{v}\,\mathrm{d}v = \int_0^{\boldsymbol{p}_D(0)} D(v)\,\mathrm{d}v,$$

where

$$\widehat{D}(v) := \begin{cases} 1, & \text{if } v = 0 \\ \frac{\pi_0}{\hat{v}}, & \text{if } v \in (0, \hat{v}] \\ \frac{\pi_0}{v}, & \text{if } v \in (\hat{v}, \boldsymbol{p}_D(0)] \\ D(v), & \text{if } v \in (\boldsymbol{p}_D(0), \infty) \end{cases}.$$

Therefore, we have $\widehat{D} \in \mathcal{D}_s$, Moreover, since $\pi_0 \geq vD(v)$ for all $v \in [0,1]$, we must have $\max_v v\widehat{D}(v) = \pi_0$ and that $\boldsymbol{p}_{\widehat{D}}(0) = \hat{v}$. This in turn implies that

$$\begin{aligned}\int_{\boldsymbol{p}_{\widehat{D}}(0)}^1 \widehat{D}(v)\,\mathrm{d}v = \int_{\hat{v}}^1 D(v)\,\mathrm{d}v =& s - \int_0^{\hat{v}} \widehat{D}(v)\,\mathrm{d}v \\ =& s - \pi_0 \\ =& s - D(\boldsymbol{p}_D(0)) \\ >& \int_{\boldsymbol{p}_D(0)}^1 D(v)\,\mathrm{d}v.\end{aligned}$$

Furthermore, for all $c > 0$, since $v \mapsto (v-c)\pi_0/v$ is increasing,

$$(v-c)\widehat{D}(v) \leq (\boldsymbol{p}_D(0) - c)\frac{\pi_0}{\boldsymbol{p}_D(0)} = (\boldsymbol{p}_D(0) - c)D(\boldsymbol{p}_D(0)) \leq (\boldsymbol{p}_D(c) - c)D(\boldsymbol{p}_D(c)),$$

where the last inequality holds if and only if $\boldsymbol{p}_D(0) < \boldsymbol{p}_D(c)$. As a result, $\boldsymbol{p}_D(c) = \boldsymbol{p}_{\widehat{D}}(c)$ for all $c \geq 0$. Together, we have $\Pi(\widehat{D}|\gamma) \geq \Pi(D|\gamma)$ and $\Sigma(\widehat{D}|\gamma) \geq \Sigma(D|\gamma)$ and

$$\int_{\boldsymbol{p}_{\widehat{D}}(0)}^{1} D(v)\,\mathrm{d}v + \boldsymbol{p}_{\widehat{D}}(0)\widehat{D}(\boldsymbol{p}_{\widehat{D}}(0)) = s.$$

Thus, by showing that there exists $\widehat{D} \in \mathcal{D}_s$ such that $\alpha\Pi(\widehat{D}|\gamma) + \beta\Sigma(\widehat{D}|\gamma) > \alpha\Pi(D/\gamma) + \beta\Sigma(D|\gamma)$ for any non affine-unit-elastic D with

$$\boldsymbol{p}_D(0)D(\boldsymbol{p}_D(0)) + \int_{\boldsymbol{p}_D(0)}^{1} D(v)\,\mathrm{d}v = s,$$

it then immediately follows that $\alpha\Pi(\widehat{D}|\gamma) + \beta\Sigma(\widehat{D}|\gamma) > \alpha\Pi(D/\gamma) + \beta\Sigma(D|\gamma)$ for all D with $D(\boldsymbol{p}_D(0)) < 1$ and

$$\boldsymbol{p}_D(0)D(\boldsymbol{p}_D(0)) + \int_{\boldsymbol{p}_D(0)}^{1} D(v)\,\mathrm{d}v < s$$

as well.

For notational convenience, we now transform the coordinates to write integrals in the value space. Specifically, notice that

$$\Sigma(D|\gamma) = \int_0^{\bar{c}} \left(\int_{\boldsymbol{p}_D(c)}^{1} D(v)\,\mathrm{d}v \right) \gamma(\mathrm{d}c) = \int_0^1 \gamma(\boldsymbol{p}_D^{-1}(v))D(v)\,\mathrm{d}v$$

for any market demand D and for any $\gamma \in \Gamma$, where

$$\boldsymbol{p}_D^{-1}(v) := \inf\{c \geq 0 | \boldsymbol{p}_D(c) \geq v\}.$$

Now consider two cases separately.

Case 1:

$$\boldsymbol{p}_D(0)D(\boldsymbol{p}_D(0)) + \int_{\boldsymbol{p}_D(0)}^{1} D(v)\,\mathrm{d}v = s. \tag{10}$$

In this case, since $D \in \mathcal{D}_s$,

$$\int_0^{\boldsymbol{p}_D(0)} (D(v) - D(\boldsymbol{p}_D(0)))\,\mathrm{d}v = 0.$$

Since D is nondecreasing, it must be that $D(v) = D(\boldsymbol{p}_D(0))$ for all $v \in (0, \boldsymbol{p}_D(0)]$. By Lemma 1, $\boldsymbol{p}_D$ takes at least three different values on $[0,1]$, which is equivalent to saying that $\boldsymbol{p}_D^{-1}$ takes at least three values on $[0,1]$.

Suppose first that there exists some $v_0 \in (0,1)$ and a sequence $\{v_n\}$ such that $v_n < v_{n+1} < c_0$ and $\boldsymbol{p}_D^{-1}(v_n) < \boldsymbol{p}_D^{-1}(v_{n+1}) < \boldsymbol{p}_D(v_0)$ for all $n \in \mathbb{N}$ and that $\{v_n\} \uparrow v_0$. Since $\boldsymbol{p}_D^{-1}(v_{n+1}) > \boldsymbol{p}_D^{-1}(v_n)$ for all $n \in \mathbb{N}$ and $\boldsymbol{p}_D^{-1}$ is nondecreasing, we may take such $\{v_n\}$ so that $\boldsymbol{p}_D^{-1}(v_n^+) > \boldsymbol{p}_D^{-1}(v')$ for all $v \in [\boldsymbol{p}_D(0), v_n)$ for all $n \in \mathbb{N}$. Let $k := \boldsymbol{p}_D^{-1}(v_0)$ and $\zeta_0 := (v_0 - k)D(v_0)$. Moreover, for each $n \in \mathbb{N}$, let $\zeta_n := (v_n - \boldsymbol{p}_D^{-1}(v_n^+))D(v_n)$. Since $\boldsymbol{p}_D^{-1}(v_n^+) > \boldsymbol{p}_D^{-1}(v')$ for all $v' \in [\boldsymbol{p}_D(0), v_n)$ and since $\boldsymbol{p}_D^{-1}(v_{n+1}) > \boldsymbol{p}_D^{-1}(v_n)$, we must have

$$(v - \boldsymbol{p}_D^{-1}(v_n))D(v) \leq (v_n - \boldsymbol{p}_D^{-1}(v_n))D(v_n),$$

and

$$(v - k)D(v) \leq (v_0 - k)D(v_0),$$

for all $v \in [v_n, v_0]$, with strict inequality at some $v \in (v_n, v_0)$. Thus,

$$D(v) \leq \min\left\{\frac{\zeta_0}{v-k}, \frac{\zeta_n}{v - \boldsymbol{p}_D^{-1}(v_n)}\right\}, \tag{11}$$

for all $v \in [v_n, v_0]$ with strict inequality at some $v \in (v_n, v_0)$. Thus, there exists $\bar{n} \in \mathbb{N}$ such that whenever $n > \bar{n}$, there exists $\underline{v}_n \in (\boldsymbol{p}_D(0), v_n)$ and $\hat{v}_n \in (v_n, v_0)$ such that

$$\int_0^{v_0} D(v)\,\mathrm{d}v = \underline{v}_n D(\underline{v}_n) + \int_{\underline{v}_n}^{v_n} D(v)\,\mathrm{d}v + \int_{v_n}^{v_0} \min\left\{\frac{\zeta_0}{v-k}, \frac{\zeta_n}{x - \boldsymbol{p}_D^{-1}(v_n)}\right\}\mathrm{d}v \tag{12}$$

$$\frac{\zeta_0}{\hat{v}_n - k} = \frac{\zeta_n}{\hat{v}_n - \boldsymbol{p}_D^{-1}(v_n)}. \tag{13}$$

As such, for any $n > \bar{n}$, define

$$\widehat{D}^{v_n}(v) := \begin{cases} D(v), & \text{if } v \in [v_0, 1] \\ \frac{\zeta_0}{v-k}, & \text{if } v \in [\hat{v}_n, v_0) \\ \frac{\zeta_n}{v - \boldsymbol{p}_D^{-1}(v_n)}, & \text{if } v \in [v_n, \hat{v}_n) \\ D(v), & \text{if } v \in [\underline{v}_n, v_n) \\ D(\underline{v}_n), & \text{if } v \in [0, \underline{v}_n) \end{cases},$$

where $\hat{v}_n$ and $\underline{v}_n$ are uniquely defined by (12) and (13). Notice that by (11), $\hat{v}_n < \hat{v}_{n+1}$ for all $n \in \mathbb{N}$, $\{\hat{v}_n\} \uparrow v_0$ and $\underline{v}_n > \underline{v}_{n+1}$ for all $n \in \mathbb{N}$, $\{\underline{v}_n\} \downarrow r$. Moreover, $\widehat{D}^{v_n} \in \mathcal{D}_s$ for all $n \in \mathbb{N}$.

By construction, for all $v \in [0, 1]$ and for any $n > \bar{n}$,

$$\boldsymbol{p}_{\widehat{D}^{v_n}}^{-1}(v) = \begin{cases} \boldsymbol{p}_D^{-1}(v), & \text{if } v \in [v_0, 1] \\ k, & \text{if } v \in [\hat{v}_n, v_0) \\ \boldsymbol{p}_D^{-1}(v_n), & \text{if } v \in [v_n, \hat{v}_n) \\ \boldsymbol{p}_D^{-1}(v), & \text{if } v \in [\underline{v}_n, v_n) \\ 0, & \text{if } v \in [0, \underline{v}_n) \end{cases}.$$

As a result, the difference in consumer surplus between under $\widehat{D}^{v_n}$ and under D is

$$\begin{aligned} &\int_0^1 \gamma(\boldsymbol{p}_{\widehat{D}^{x_n}}^{-1}(v))\widehat{D}^{v_n}(v)\,\mathrm{d}v - \int_0^1 \gamma(\boldsymbol{p}_D^{-1}(v)D(v)\,\mathrm{d}v \\ =&\gamma(k)(\boldsymbol{p}_D(0)D(\boldsymbol{p}_D(0)) - \underline{v}_n D(\underline{v}_n)) + \int_{\boldsymbol{p}_D(0)}^{\underline{v}_n} (\gamma(k) - \gamma(\boldsymbol{p}_D^{-1}(v)))D(v))\,\mathrm{d}v \\ &- \int_{v_n}^{\hat{v}_n} \big((\gamma(k) - \gamma(\boldsymbol{p}_D^{-1}(v_n)))\widehat{D}^{v_n}(v)\big)\,\mathrm{d}v \\ &+ \int_{v_n}^{v_0} \big(\gamma(k) - \gamma(\boldsymbol{p}_D{-}1(v)))D(x)\big)\,\mathrm{d}v. \end{aligned}$$

Notice that by definition, $\boldsymbol{p}_D(0)D(\boldsymbol{p}_D(0)) \geq \underline{v}_n D(\underline{v}_n)$ for all $n > \in \mathbb{N}$.

For any $\underline{v} \in (\boldsymbol{p}_D(0), v_0)$, let

$$\begin{aligned} \Delta(\underline{v}) := \int_{\boldsymbol{p}_D(0)}^{\underline{v}} \big(\gamma(k) - \gamma(\boldsymbol{p}_D^{-1}(v))\big)D(v)\,\mathrm{d}v - &\int_{\bar{v}(\underline{v})}^{\hat{v}(\underline{v})} \big(\gamma(k) - \gamma(\boldsymbol{p}_D^{-1}(\bar{v}(\underline{v}))\big)\widehat{D}^{\underline{v}}(v)\,\mathrm{d}v \\ + &\int_{\bar{v}(\underline{v})}^{v_0} \big(\gamma(k) - \gamma(\boldsymbol{p}_D^{-1}(v))\big)D(v)\,\mathrm{d}v, \end{aligned} \tag{14}$$

where $\hat{v}(\underline{v})$ and $\bar{v}(\underline{v})$ are uniquely defined by

$$\int_0^{v_0} D(x)\,\mathrm{d}v = \underline{v}D(\underline{v}) + \int_{\underline{v}}^{\bar{v}(\underline{v})} D(v)\,\mathrm{d}v + \int_{\bar{v}(\underline{v})}^{v_0} \min\left\{\frac{\zeta_0}{v-k}, \frac{\zeta_{\bar{v}(\underline{v})}}{v-\boldsymbol{p}_D^{-1}(\bar{v}(\underline{v}))}\right\}\mathrm{d}v \tag{15}$$

$$\frac{\zeta_0}{\hat{v}(\underline{v})-k} = \frac{\zeta_{\bar{v}(\underline{v})}}{\hat{v}(\underline{v})-\boldsymbol{p}_D^{-1}(\bar{v}(\underline{v}))}, \tag{16}$$

and $\widehat{D}^{\underline{v}}$ is defined by

$$\widehat{D}^{\underline{v}}(v) := \begin{cases} D(v), & \text{if } v \in [v_0, 1] \\ \frac{\zeta_0}{v-k}, & \text{if } v \in [\hat{v}(\underline{v}), x_0) \\ \frac{\zeta_{\bar{v}(\underline{v})}}{v-\boldsymbol{p}_D^{-1}(\bar{v}(\underline{v}))}, & \text{if } x \in [\bar{v}(\underline{v}), \hat{v}(\underline{v})) \\ D(v), & \text{if } v \in [\underline{v}, \bar{v}(\underline{v})) \\ D(\underline{v}), & \text{if } v \in [0, \underline{v}) \end{cases},$$

where $\zeta_v := (v-\boldsymbol{p}_D^{-1}(v))D(v)$ for all $v \in [\boldsymbol{p}_D(0), 1]$.

By (12), (13), (15) and (16), for any $n > \bar{n}$, we have $v_n = \bar{v}(\underline{v}_n)$ and $\hat{v}_n = \hat{v}(\underline{v}_n)$. Also, $\hat{v}$ and $\bar{v}$ are decreasing in $\underline{v}$ and $\lim_{\underline{v}\to\boldsymbol{p}_D(0)} \bar{v}(\underline{v}) = \lim_{\underline{v}\to\boldsymbol{p}_D(0)} \hat{v}(\underline{v}) = v_0$, $\lim_{\underline{v}\to\boldsymbol{p}_D(0)} \boldsymbol{p}_D^{-1}(\bar{v}(\underline{v})) = \lim_{n\to\infty} \boldsymbol{p}_D^{-1}(v_n) = k$.

Furthermore, since $\boldsymbol{p}_D^{-1}$ and γ are nondecreasing, by (15) and (16), $\bar{v}$ and $\hat{v}$ are differentiable Lebesgue-almost everywhere and therefore Δ is differentiable Lebesgue-almost everywhere. Thus, for Lebesgue almost all $\underline{v}$,

$$\begin{aligned}\Delta'(\underline{v}) = (\gamma(k)-\gamma(\boldsymbol{p}_D^{-1}(\underline{v})))D(\underline{v}) - \hat{v}'(\underline{v})(\gamma(k)-\gamma(\boldsymbol{p}_D^{-1}(\bar{v}(\underline{v}))))\widehat{D}^{\underline{v}}(\hat{v}(\underline{v})) \\ - \int_{\bar{v}(\underline{v})}^{\hat{v}(\underline{v})} \frac{\partial}{\partial \underline{v}}(\gamma(k)-\gamma(\boldsymbol{p}_D^{-1}(\bar{v}(\underline{v}))))\widehat{D}^{\underline{v}}(v)\,\mathrm{d}v.\end{aligned}$$

As $\lim_{\underline{v}\to\boldsymbol{p}_D(0)} \hat{v}(\underline{v}) = \lim_{\underline{v}\to\boldsymbol{p}_D(0)} \bar{v}(\underline{v}) = v_0$, $\lim_{\underline{v}\to\boldsymbol{p}_D(0)} \boldsymbol{p}_D^{-1}(\bar{v}(\underline{v})) = k$ and since $\hat{v}$ is decreasing, there exists $\delta > 0$ such that for $\underline{v}$ sufficiently close to $\boldsymbol{p}_D(0)$, whenever Δ is differentiable at $\underline{v}$,

$$\Delta'(\underline{v}) > (\gamma(k)-\gamma(\boldsymbol{p}_D^{-1}(\underline{v})))D(\underline{v}) - \delta > 0, \tag{17}$$

where the second inequality follows from the hypothesis that γ is strictly increasing. Therefore, since Δ is continuous in $\underline{v}$ and $\Delta'(\underline{v}) > 0$ for $\underline{v}$ sufficiently close to $\boldsymbol{p}_D(0)$, there exists $\hat{v}$ such that $\Delta(\underline{v}) > 0$ for all $\underline{v} \in (\boldsymbol{p}_D(0), \hat{v})$.

As such, for n sufficiently large so that $\underline{v}_n \in (\boldsymbol{p}_D(0), \hat{v})$,

$$\begin{aligned}0 < \Delta(\underline{v}_n) = &\int_{\boldsymbol{p}_D(0)}^{\underline{v}_n} (\gamma(k)-\gamma(\boldsymbol{p}_D^{-1}(v)))D(v))\,\mathrm{d}v \\ &- \int_{v_n}^{\hat{v}_n} \big((\gamma(k)-\gamma(\boldsymbol{p}_D^{-1}(v_n)))\widehat{D}^{v_n}(v)\big)\,\mathrm{d}v \\ &+ \int_{v_n}^{v_0} \big(\gamma(k)-\gamma(\boldsymbol{p}_D^{-1}(v)))D(v)\big)\,\mathrm{d}v.\end{aligned}$$

Together, for n sufficiently large,

$$\Sigma(\widehat{D}^{v_n}|\gamma) = \int_0^1 \gamma(\boldsymbol{p}_{\widehat{D}^{v_n}}^{-1}(v))\widehat{D}^{v_n}(v)\,\mathrm{d}v > \int_0^1 \gamma(\boldsymbol{p}_D^{-1}(v))D(v)\,\mathrm{d}v = \Sigma(D|\gamma),$$

as desired.

As for the monopolist's profit, notice that for each $n \in \mathbb{N}$, any $c \in [0,1]\backslash(\boldsymbol{p}_D^{-1}(v_n), k) \cup (0, \boldsymbol{p}_D^{-1}(\underline{v}_n|R))$, $\max_p(p-c)D(p) = \max_p(p-c)\widehat{D}^{v_n}$. Therefore,

$$\begin{aligned}&\Pi(D|\gamma) - \Pi(\widehat{D}^{v_n}|\gamma)\\ =& \int_0^{\boldsymbol{p}_D^{-1}(\underline{v}_n)} [\max_p(p-c)D(p) - \max_p(p-c)\widehat{D}^{\underline{v}_n}(p)]\gamma(\mathrm{d}c) + \int_{\boldsymbol{p}_D^{-1}(v_n)}^k [\max_p(p-c)D(p) - \max_p(p-c)\widehat{D}^{v_n}]\gamma(\mathrm{d}c)\end{aligned}$$

Let

$$\Lambda(\underline{v}) := \int_0^{\boldsymbol{p}_D^{-1}(\underline{v})} [\max_p(p-c)D(p) - \max_p(p-c)\widehat{D}^{\underline{v}}(p)]\gamma(\mathrm{d}c) + \int_{\boldsymbol{p}_D^{-1}(v_n)}^k [\max_p(p-c)D(p) - \max_p(p-c)\widehat{D}^{v}]\gamma(\mathrm{d}c) \tag{18}$$

Notice that $\Lambda(\boldsymbol{p}_D(0)) = 0$. Furthermore, since $\boldsymbol{p}_D^{-1}(\underline{v}) \to 0$ and $\boldsymbol{p}_D^{-1}(\bar{v}(\underline{v})) \to k$ as $\underline{v} \to \boldsymbol{p}_D(0)$, and since $\max_p pD(p) = \max_p p\widehat{D}^{\underline{v}}(p)$ and $\max_p(p-k)D(p) = \max_p(p-k)\widehat{D}^{\underline{v}}(p)$ for all $\underline{v}$, we have

$$\Lambda'(\boldsymbol{p}_D(0)) = \gamma'(0)(\max_p pD(p) - \max_p p\widehat{D}^{\underline{v}}(p)) + \gamma'(k)(\max_p(p-k)D(p) - \max_p(p-k)\widehat{D}^{\underline{v}}(p)) = 0. \tag{19}$$

Together, for $n \in \mathbb{N}$ large enough,

$$\alpha\Pi(D|\gamma) + \beta\Sigma(D|\gamma) < \alpha\Pi(\widehat{D}^{v_n}|\gamma) + \beta\Sigma(\widehat{D}^{v_n}|\gamma),$$

as desired.

Secondly, consider the case where for any $v \in (\boldsymbol{p}_D(0), 1)$ and for any sequence $\{v_n\}$ such that $\{v_n\} \uparrow v$, there exists $\bar{n} \in \mathbb{N}$ such that $\boldsymbol{p}_D^{-1}(v_n) = \boldsymbol{p}_D^{-1}(v)$ for all $n > \bar{n}$, then for any $v \in (\boldsymbol{p}_D(0), 1)$, there exists $\delta > 0$ such that $\boldsymbol{p}_D^{-1}(v'|R) = \boldsymbol{p}_D^{-1}(v)$ for all $v' \in (v-\delta, v)$. Let $\delta_v := \sup\{\delta > 0 | \boldsymbol{p}_D^{-1}(v') = \boldsymbol{p}_D^{-1}(v), \forall v' \in (v-\delta, v)\}$. Then $\delta_v > 0$ for all $v \in (\boldsymbol{p}_D(0), 1)$. Moreover, for any $v, v' \in (\boldsymbol{p}_D(0), 1)$, if $\boldsymbol{p}_D^{-1}(v) \neq \boldsymbol{p}_D^{-1}(v')$, then it must be that $(v - \delta_v, v) \cap (v' - \delta_{v'}, v') = \emptyset$. Therefore, $\{\boldsymbol{p}_D^{-1}(v)\}_{v \in (\boldsymbol{p}_D(0),1)}$ is at most countable. Since $\boldsymbol{p}_D^{-1}$is nondecreasing, it must be a step function. Let $\bar{v} := \sup\{v \in [0,1] | D(v) > 0\}$ and consider first the case when for any $\delta > 0$, there exists $v, v' \in (\bar{v} - \delta, \bar{v})$ with $v < v'$; such that $\boldsymbol{p}_D^{-1}(v) < \boldsymbol{p}_D^{-1}(v') < \boldsymbol{p}_D^{-1}(\bar{v})$. Since $\boldsymbol{p}_D^{-1}$ is a step function, $\boldsymbol{p}_D^{-1}$ has countably infinitely many jumps and therefore we may represent $\boldsymbol{p}_D^{-1}$ as

$$\boldsymbol{p}_D^{-1}(v) = \sum_{n=1}^{\infty} c_n \mathbf{1}\{v \in (a_n, b_n)\}, \forall v \in [\boldsymbol{p}_D(0), 1] \backslash [\{a_n\}_{n=1}^{\infty} \cup \{b_n\}_{n=1}^{\infty}].$$

for some $\{a_n\}$, $\{b_n\}$ with $a_n < b_n$ for all $n \in \mathbb{N}$ and some $\{c_n\}_{n=1}^{\infty}$ with $c_n > 0$ for all $n \in \mathbb{N}$, $n \geq 2$. Since for any $\delta > 0$, there exists $v, v' \in (\bar{v} - \delta, \bar{v})$, $v < v'$, such that $\boldsymbol{p}_D^{-1}(v) < \boldsymbol{p}_D^{-1}(v') < \boldsymbol{p}_D^{-1}(\bar{v})$, there exists a sequence $\{v_j\}$ such that $v_j \in \{a_n\}_{n=1}^{\infty} \cup \{a_n\}_{n=1}^{\infty}$, $v_j < v_{j+1}$, $c_j := \boldsymbol{p}_D^{-1}(v_j^+) = \boldsymbol{p}_D^{-1}(v_{j+1}) < \boldsymbol{p}_D^{-1}(v_{j+1}^+) =: v_{j+1}$ for all $j \in \mathbb{N}$, and that $\{v_j\} \uparrow \bar{v}$ and $\{c_j\} \uparrow \boldsymbol{p}_D^{-1}(\bar{v})$.

Since $\boldsymbol{p}_D^{-1}$ is a left-continuous step function and $\boldsymbol{p}_D^{-1}(v) = \bar{v}$ for all $v \in [b, 1]$, we must have $\boldsymbol{p}_D^{-1}(\bar{v}) < \bar{v}$ and $D(\bar{v}) > 0$. As such, let $\bar{\zeta} := (\bar{v} - \boldsymbol{p}_D^{-1}(\bar{v}))D(\bar{v})$. Then $\bar{\zeta} > 0$. Also, for each $j \in \mathbb{N}$, let $\zeta_j := (v_j - c_j)D(v_j)$. Then, since $\psi_{j-1} < \psi_j < \psi_{j+1}$, we have

$$(v - c_j)D(v) < (v_j - c_j)D(v_j), \text{ and } (v - \boldsymbol{p}_D^{-1}(\bar{v}))D(\bar{v}) < (\bar{v} - \boldsymbol{p}_D^{-1}(\bar{v}))D(\bar{v})$$

for all $v \in (v_j, \bar{v})$, and hence

$$D(x) < \min\left\{\frac{\bar{\zeta}}{x - \boldsymbol{p}_D^{-1}(\bar{v})}, \frac{\zeta_j}{v - c_j}\right\},$$

for all $v \in (v_j, \bar{v})$.

As such, there exists a sequence $\{\underline{v}_j\}$ such that $\{\underline{v}_j\} \downarrow \boldsymbol{p}_D(0)$ such that $\widehat{D}^j \in \mathcal{D}_s$ for j large enough, where $\widehat{D}^j$ is defined as:

$$\widehat{D}^j(v) := \begin{cases} D(v), & \text{if } v \in [\bar{v}, 1] \\ \frac{\bar{\zeta}}{v - \boldsymbol{p}_D^{-1}(\bar{v})}, & \text{if } v \in (\hat{v}_j, \bar{v}] \\ \frac{\zeta_j}{v - \psi_j}, & \text{if } v \in (\bar{v}_j, \hat{v}_j] \\ D(v), & \text{if } x \in (\underline{v}_j, \bar{v}_j] \\ D(\underline{v}_j), & \text{if } v \in (0, \underline{v}_j] \end{cases},$$

where $v_j < \hat{v}_j < \bar{v}_j < \bar{v}$ are uniquely defined by

$$\int_0^1 \widehat{D}^j(v)\,\mathrm{d}v = \int_0^1 D(v)\,\mathrm{d}v$$
$$\frac{\bar{\zeta}}{\hat{v}_j - \boldsymbol{p}_D^{-1}(\bar{v})} = \frac{\zeta_j}{\hat{v}_j - c_j},$$

Similar to the previous case, for each $j \in \mathbb{N}$ such that $\widehat{D}^j \in \mathcal{D}_s$, the deviation gain from R to $\widehat{D}^j$ is

$$\gamma(\boldsymbol{p}_D^{-1}(\bar{v}))(\boldsymbol{p}_D(0)D(\boldsymbol{p}_D(0)) - \underline{v}_j D(\underline{v}_j)) + \int_{\boldsymbol{p}_D(0)}^{\underline{v}_j} (\gamma(\boldsymbol{p}_D^{-1}(\bar{v})) - \gamma(\boldsymbol{p}_D^{-1}(v)))D(v)\,\mathrm{d}v$$
$$- \int_{\bar{v}_j}^{\hat{v}_j} \big(\gamma(\boldsymbol{p}_D^{-1}(\bar{v})) - \gamma(c_j)\big)\widehat{D}^j(v)\,\mathrm{d}v$$
$$+ \int_{\bar{v}_j}^{v_j} \big(\gamma(\boldsymbol{p}_D^{-1}(\bar{v})) - \gamma(\boldsymbol{p}_D^{-1}(v))\big)D(v)\,\mathrm{d}v.$$

By definition, $\boldsymbol{p}_D^{-1}(0)D(\boldsymbol{p}_D(0)) \geq \underline{v}_j D(\underline{v}_j)$ for all $j \in \mathbb{N}$. Also, as shown in the previous case, from (14) and (17), with $v_0 = b$ and $k = \boldsymbol{p}_D^{-1}(\bar{v})$, there exists $\delta > 0$ such that for $\underline{v}$ sufficiently close to $\boldsymbol{p}_D(0)$, $\Delta'(\underline{v}) > (\gamma(k) - \boldsymbol{p}_D^{-1}(0^+))D(\boldsymbol{p}_D(0)^+) - \delta > 0$ since $c_j > 0$ for all $j \in \mathbb{N}$ and γ is strictly increasing. Thus, since by (15) and (16), $\hat{v}_j = \hat{v}(\underline{v}_j)$ and $\bar{v}_j = \bar{v}(\underline{v}_j)$, and since $\{\underline{v}_j\} \downarrow \boldsymbol{p}_D(0)$, for j sufficiently large,

$$0 < \Delta(\underline{v}_j) = \int_{\boldsymbol{p}_D(0)}^{\underline{v}_j} (\gamma(\boldsymbol{p}_D^{-1}(\bar{v})) - \gamma(\boldsymbol{p}_D^{-1}(v))D(v)\,\mathrm{d}v$$
$$- \int_{\bar{v}_j}^{\hat{v}_j} \big(\gamma(\boldsymbol{p}_D^{-1}(\bar{v})) - \gamma(c_j)\big)\widehat{D}^j(v)\,\mathrm{d}v$$
$$+ \int_{\bar{v}_j}^{v_j} \big(\gamma(\boldsymbol{p}_D^{-1}(\bar{v})) - \gamma(\boldsymbol{p}_D^{-1}(v))\big)D(v)\,\mathrm{d}v.$$

Together, for j large enough, $\widehat{D}^j \in \mathcal{D}_s$ and

$$\Sigma(\widehat{D}^j|\gamma) = \int_0^1 \gamma(\boldsymbol{p}_{\widehat{D}^j}^{-1}(v))\widehat{D}^j(v))\,\mathrm{d}v > \int_0^1 \gamma(\boldsymbol{p}_D^{-1}(v))D(v)\,\mathrm{d}v.$$

In the mean time, the difference in monopolist's profit between D and $\widehat{D}^j$ is

$$\int_0^1 \max_p (p - c)D(p)\gamma(\mathrm{d}c) - \int_0^1 \max_p (p - c)\widehat{D}_j(p)\gamma(\mathrm{d}c).$$

Similar to the previous case, for all $j \in \mathbb{N}$, $\max_p (p-c)D(p) = \max_p (p-c)\widehat{D}_j(p)$ for all $c \in [0,1] \backslash (c_j, \boldsymbol{p}_D^{-1}(\bar{v})) \cup (0, \boldsymbol{p}_D^{-1}(\underline{v}_j))$. Thus, by (18) and (19), $\Lambda'(\boldsymbol{p}_D(0)) = 0$ implies that the first order difference of the monopolist's profit is zero. Together, there exists $\widehat{D}^j \in \mathcal{D}_s$ such that

$$\alpha \Pi(\widehat{D}^j|\gamma) + \beta \Sigma(\widehat{D}^j|\gamma) > \alpha \Pi(D|\gamma) + \beta \Sigma(D|\gamma),$$

as desired.

Lastly, consider the case where there exists $\delta > 0$ such that for any $v, v' \in (\bar{v}-\delta, \bar{v})$, $\boldsymbol{p}_D^{-1}(v) = \boldsymbol{p}_D^{-1}(v') = \boldsymbol{p}_D^{-1}(\bar{v})$. Let $v_* := \inf\{v \in [r,1) | \boldsymbol{p}_D^{-1}(v) = \boldsymbol{p}_D^{-1}(v'), \forall v, v' \in (v, \bar{v})\}$ and let $c_* := \boldsymbol{p}_D^{-1}(v_*^+)$. Then $v_* < \bar{v}$ and $c_* > \boldsymbol{p}_D^{-1}(\boldsymbol{p}_D(0)^+)$ since $\boldsymbol{p}_D^{-1}$ not constant on $[\boldsymbol{p}_D(0), \bar{v})$.

By definition, since $\boldsymbol{p}_D^{-1}(v) = c_*$ for all $v \in (v_*, \bar{v})$, we must have $(v_*, \bar{v}) \subseteq \operatorname{argmax}_p (p-c)D(p)$ and hence

$$(v - c_*)D(v) = \zeta \iff D(v) = \frac{\zeta}{v - c_*}, \forall v \in (v_*, \bar{v}),$$

for some $\zeta > 0$. Now fix any $\tilde{v} \in (v_*, \bar{v})$ and notice that for any $\tilde{c} \in (\boldsymbol{p}_D^{-1}(v_*), c_*)$ and any $\hat{c} \in (c_*, \bar{v})$, let $\zeta(\hat{c}) := (\tilde{v} - \hat{c})D(\tilde{v})$ and $\zeta(\tilde{c}) := (v_* - \tilde{c})D(v_*)$, we must have

$$D(v) < \min\left\{\frac{\zeta(\hat{c})}{v - \hat{c}}, \frac{\zeta(\tilde{c})}{v - \tilde{c}}\right\},$$

for all $v \in (v_*, \tilde{v})$. Thus, for any $\underline{v} > \boldsymbol{p}_D(0)$ that is close enough to $\boldsymbol{p}_D(0)$, there exists $\tilde{c}(\underline{v}) \in (\boldsymbol{p}_D^{-1}(v_*), c_*)$,$\hat{c}(\underline{v}) \in (c_*, \bar{v})$ and $\hat{v}(\underline{v}) \in (\underline{v}, \bar{v})$ such that $\lim_{\underline{v} \downarrow \boldsymbol{p}_D(0)} \tilde{c}(\underline{v}) = \lim_{\underline{v} \downarrow \boldsymbol{p}_D(0)} \hat{c}(\underline{v}) = c_*$,

$$\frac{\zeta(\hat{c}(\underline{v}))}{\hat{v}(\underline{v}) - \hat{c}(\underline{v})} = \frac{\zeta(\tilde{c}(\underline{v}))}{\hat{v}(\underline{v}) - \tilde{v}(\underline{v})}$$

and

$$\int_0^{\tilde{v}} D(v)\,\mathrm{d}v = \int_0^{\tilde{v}} \widehat{D}^{\underline{v}}(v)\,\mathrm{d}v,$$

where

$$\widehat{D}^{\underline{v}}(v) := \begin{cases} D(v), & \text{if } v \in [\tilde{v}, 1] \\ \frac{\zeta(\hat{c}(\underline{v}))}{v - \hat{c}(\underline{v})}, & \text{if } v \in (\hat{v}(\underline{x}), \tilde{v}\underline{v}] \\ \frac{\zeta(\tilde{v}(\underline{v}))}{v - \tilde{v}(\underline{v})}, & \text{if } v \in (v_*, \hat{v}(\underline{v})) \\ D(v), & \text{if } v \in (\underline{v}, v_*] \\ D(\underline{v}), & \text{if } v \in (0, \underline{v}] \end{cases}$$

and thus $\widehat{D}^{\underline{v}} \in \mathcal{D}_s$. Moreover, such $\tilde{c}(\cdot)$ and $\hat{c}(\cdot)$ can be selected so that $\hat{v}(\underline{v})$ is decreasing in $\underline{v}$ and $\lim_{\underline{v} \downarrow \boldsymbol{p}_D(0)} \hat{v}(\underline{v}) = \tilde{v}$.

Notice that for any such $\widehat{D}^{\underline{v}}$,

$$\boldsymbol{p}_{\widehat{D}^{\underline{v}}}^{-1}(v) = \begin{cases} 0, & \text{if } v \in [0, \underline{v}] \\ \boldsymbol{p}_D^{-1}(v), & \text{if } v \in (\underline{v}, v_*) \\ \tilde{c}(\underline{v}), & \text{if } v \in [v_*, \hat{v}(\underline{v})) \\ \tilde{k}(\underline{v}), & \text{if } v \in [\hat{v}(\underline{v}), \bar{v}) \\ \bar{v}, & \text{if } x \in [\bar{v}, 1]. \end{cases},$$

for some $\tilde{k}(\underline{v}) \in (c_*, k(\underline{v}))$. As such, for any $\underline{v}$ such that $\widehat{D}^{\underline{v}} \in \mathcal{D}_s$, the deviation gain from D to $\widehat{D}^{\underline{v}}$ is

$$\begin{aligned}
&\int_0^1 \gamma(\boldsymbol{p}_{\widehat{D}^{\underline{v}}}(v)\widehat{D}^{\underline{v}}(v)\,\mathrm{d}v - \int_0^1 \gamma(\boldsymbol{p}_D^{-1}(v))D(v))\,\mathrm{d}v \\
&> \gamma(\tilde{k}(\underline{v}))(\boldsymbol{p}_D(0)D(\boldsymbol{p}_D(0)) - \underline{v}D(\underline{v})) + \int_{\boldsymbol{p}_D(0)}^{\underline{v}} (\gamma(\tilde{k}(\underline{v})) - \gamma(\boldsymbol{p}_D^{-1}(v))D(v)\,\mathrm{d}v \\
&- \int_{v_*}^{\hat{v}(\underline{v})} (\gamma(\tilde{k}(\underline{v})) - \gamma(\tilde{c}(\underline{v})))\widehat{D}^{\underline{v}}(v)\,\mathrm{d}v,
\end{aligned}$$

where the strict inequality follows from $\tilde{k}(\underline{v}) > c_*$ and that γ is strictly increasing.

Again, by definition, $\boldsymbol{p}_D(0)D(\boldsymbol{p}_D(0)) \geq \underline{v}D(\underline{v})$ for all $\underline{v} \geq \boldsymbol{p}_D(0)$. Also, let

$$\Delta(\underline{v}) := \int_{\boldsymbol{p}_D(0)}^{\underline{v}} (\gamma(\tilde{k}(\underline{v})) - \gamma(\boldsymbol{p}_D^{-1}(v)))D(v)\,\mathrm{d}v - \int_{v_*}^{\hat{v}(\underline{v})} (\gamma(\tilde{k}(\underline{v})) - \gamma(\tilde{c}(\underline{v})))\widehat{D}^{\underline{v}}(v)\,\mathrm{d}v.$$

Then Δ is differentiable Lebesgue-almost everywhere and the derivative converges to

$$(\gamma(c_*) - \gamma(\boldsymbol{p}_D^{-1}(\boldsymbol{p}_D(0)^+))D(\boldsymbol{p}_D(0)) - \lim_{\hat{c},\tilde{c}\to c_*,\,\hat{c}>\tilde{c}} (\gamma(\hat{c}) - \gamma(\tilde{c})) \cdot \lim_{\underline{v}\downarrow \boldsymbol{p}_D(0)} \hat{v}'(\underline{v}) \cdot D(\tilde{x}) > 0$$

as $\underline{v}$ approaches $\boldsymbol{p}_D(0)$, since $\hat{v}(\underline{v})$ is decreasing in $\underline{v}$, $c_* > \boldsymbol{p}_D^{-1}(\boldsymbol{p}_D(0)^+)$ and γ is strictly increasing. Thus, $\Delta(\underline{v}) > 0$ for $\underline{v}$ sufficiently close to $\boldsymbol{p}_D(0)$.

Together, for $\underline{v}$ close enough to $\boldsymbol{p}_D(0)$, there exists $\widehat{D}^{\underline{v}} \in \mathcal{D}_s$ such that

$$\Sigma(\widehat{D}^{\underline{v}}|\gamma) = \int_0^1 \gamma(\boldsymbol{p}_{\widehat{D}^{\underline{v}}}^{-1}(v)\widehat{D}^{\underline{v}}(v)\,\mathrm{d}v > \int_0^1 \gamma(\boldsymbol{p}_D^{-1}(v))D(v)\,\mathrm{d}v = \Sigma(D|\gamma).$$

Meanwhile, let

$$\Lambda(\underline{v}) := \int_0^1 \max_p (p-c)D(p)\gamma(\mathrm{d}c) - \int_0^1 \max_p (p-c)\widehat{D}^{\underline{v}}(p)\gamma(\mathrm{d}c).$$

Again, for any $\underline{v} > \boldsymbol{p}_D(0)$, for any $c \in [0,1]\backslash(0, \boldsymbol{p}_D^{-1}(\hat{v}(\underline{v}))) \cup (\tilde{c}(\underline{v}), \tilde{k}(\underline{v}))$, $\max_p(p-c)D(p) = \max_p(p-c)\widehat{D}^{\underline{v}}(p)$, which in turn, as in the previous case, implies that

$$\Lambda'(\boldsymbol{p}_D(0)) = 0.$$

Together, for $\underline{v}$ close enough to $\boldsymbol{p}_D(0)$, there exists $\widehat{D}^{\underline{v}} \in \mathcal{D}_s$ such that

$$\alpha\Pi(\widehat{D}^{\underline{v}}|\gamma) + \beta\Sigma(\widehat{D}^{\underline{v}}|\gamma) > \alpha\Pi(D|\gamma) + \beta\Sigma(D|\gamma),$$

as desired.

Case 3: $D(\boldsymbol{p}_D(0)) = 1$

Notice that the arguments in **case 2** rely only on the observations that $D(v) = D(\boldsymbol{p}_D(0))$ for all $v \in (0, \boldsymbol{p}_D(0)]$ As such, if $D(\boldsymbol{p}_D(0)^+) = 1$, then we are back to **case 2**. Meanwhile, if $D(\boldsymbol{p}_D(0)^+) < D(\boldsymbol{p}_D(0)) = 1$, we may define

$$Q(v) := \begin{cases} D(\boldsymbol{p}_D(0)), & \text{if } v \in [0, \boldsymbol{p}_D(0)] \\ D(v), & \text{if } v \in (r, 1]. \end{cases}.$$

Then since the the arguments in **case 2** do not depend on particular value of s, by the same arguments, there exists $\widehat{Q}^{\underline{v}} \in \mathcal{D}_{\bar{s}}$ such that

$$\alpha\Pi(\widehat{Q}^{\underline{v}}|\gamma) + \beta\Sigma(\widehat{Q}^{\underline{v}}|\gamma) > \alpha\Pi(Q|\gamma) + \beta\Sigma(Q|\gamma)$$

for $\underline{v} > \boldsymbol{p}_Q(0) = \boldsymbol{p}_D(0)$ small enough; and that $Q(\underline{v}) = \widehat{Q}(\underline{v}) = \widehat{Q}(0^+)$ for some $\underline{v} \in (\boldsymbol{p}_D(0), 1)$, where $\bar{s} := \int_0^1 Q(v)\,\mathrm{d}v$. Since $\underline{v} > \boldsymbol{p}_D(0)$ and $Q(\underline{v}) \leq Q(\boldsymbol{p}_D(0)) = D(\boldsymbol{p}_D(0))$, there exists $\varepsilon > 0$ such that $\underline{v}(\widehat{Q}(\underline{x}) - \varepsilon) = \boldsymbol{p}_D(0)D(\boldsymbol{p}_D(0))$ and therefore $\widehat{D} \in \mathcal{D}_s$, where

$$\widehat{D}^{\underline{v}}(v) := \begin{cases} \widehat{Q}(\underline{v}) - \varepsilon, & \text{if } x \in (0, \underline{v}] \\ \widehat{Q}^{\underline{v}}(v), & \text{if } v \in (\underline{v}, 1] \end{cases}.$$

Furthermore, by construction, $\boldsymbol{p}^{-1}_{\widehat{Q}^{\underline{v}}}(v) = \boldsymbol{p}^{-1}_{\widehat{D}^{\underline{v}}}(v)$ for all $v \in [0,1]$, $\boldsymbol{p}^{-1}_D(v) = \boldsymbol{p}^{-1}_Q(v)$ for all $v \in [0,1]$, Together, we have

$$\begin{aligned} \alpha\Pi(D|\gamma) + \beta\Sigma(D|\gamma) = \alpha\Pi(Q|\gamma) + \beta\Sigma(Q|\gamma) & \\ <& \alpha\Pi(\widehat{Q}^{\underline{v}}|\gamma) + \beta\Sigma(\widehat{Q}^{\underline{v}}|\gamma) \\ =& \alpha\Pi(\widehat{D}^{\underline{v}}|\gamma) + \beta\Sigma(\widehat{D}^{\underline{v}}|\gamma) \end{aligned}$$

for some $\widehat{D}^{\underline{v}} \in \mathcal{D}_s$, as desired. This completes the proof. ■

B Proof of Proposition 3

Proof of Proposition 3. By Theorem 1, there exists an affine-unit-elastic demand $D^{\eta}_{\pi,k} \in \mathcal{D}_s$ such that

$$\begin{aligned} W^{\alpha,\beta}(s,\gamma) =& \alpha\Pi(D^{\eta}_{\pi,k}|\gamma) + \beta\Sigma(D^{\eta}_{\pi,k}|\gamma) \\ =& \alpha\left[\int_0^k [\pi + (k-c)\eta]\gamma(\mathrm{d}c) + \int_k^{\bar{c}} (1-c)\frac{\pi}{1-k}\gamma(\mathrm{d}c)\right] + \beta[\gamma(k)(s - \pi - \eta k)] \\ =& \alpha\left[\frac{\pi}{1-k}\left(\int_0^{\bar{c}} \gamma(c)\,\mathrm{d}c + (1-\bar{c})^+\right) + \left(\eta - \frac{\pi}{1-k}\int_0^k \gamma(c)\,\mathrm{d}c\right)\right] + \beta\gamma(k)[s - \pi - \eta k)]. \end{aligned}$$

Let $\delta_1 := \alpha\pi/(1-k)$, $\delta_2 := \alpha(\eta - \pi/(1-k))$, and $\delta_3 := \beta(s - \pi - \eta k)$. Clearly $\delta_1, \delta_2, \delta_3 \geq 0$. Moreover, for any $\kappa : \mathbb{R}_+ \to \mathbb{R}$, define a function $\xi_\kappa : \mathbb{R}_+ \times [0,1]$ as follows:

$$\xi_\kappa(c|\lambda) := (1+\lambda)e^{\frac{-\delta_2 c}{\delta_3}} \int_0^c \frac{e^{\frac{\delta_2 x}{\delta_3}}(\kappa(x) - \bar{\xi})}{\delta_3}\,\mathrm{d}x,$$

for all $c \geq 0$ and for all $\lambda \in [0,1]$, where

$$\bar{\xi} := [1 + \delta_1(e^{\frac{\delta_2 \bar{c}}{\delta_3}} - 1)]^{-1} \int_0^{\bar{c}} \frac{e^{\frac{\delta_2 x}{\delta_3}}}{\delta_3}\,\mathrm{d}x > 0.$$

Then $\xi_\kappa(\cdot|0)$ solves the following differential equation:

$$\delta_1\xi_\kappa(\bar{c}|0) + \delta_2\xi_\kappa(c|0) + \delta_3\xi'_\kappa(c|0) = \kappa(c),\ \xi_\kappa(0|0) = 0.$$

Now let

$$\kappa^*(c) := \begin{cases} \bar{\kappa}, & \text{if } c \in [0, k-\varepsilon_1) \\ 1 - \frac{\bar{\kappa}+\varepsilon_2}{\varepsilon_1}(k-c), & \text{if } c \in [k-\varepsilon_1, k] \\ -\varepsilon_2, & \text{if } c \in (k, \infty) \end{cases},$$

where $\varepsilon_1 > 0$ and $\varepsilon_2 > 0$ is small enough and $\bar{\kappa} > 0$ is large enough so that $\xi_{\kappa^*}(c|0) > 0$ for all $c \in [0, \bar{c}]$. Then, for each $\lambda \in [0, 1]$, let

$$\gamma_\lambda(c) := \gamma(c) + \int_0^\lambda \frac{\partial}{\partial \lambda} \xi'_{\kappa^*}(c|y)\,\mathrm{d}y,$$

for all $\lambda \in [0, 1]$ and for all $c \in [0, \bar{c}]$. By construction, $\lambda \mapsto \gamma_\lambda(c)$ is differentiable for all $c \in [0, \bar{c}]$, $\gamma_0(c) = \gamma(c)$ for all c, and

$$\begin{aligned}
\frac{\partial}{\partial \lambda} W^{\alpha,\beta}(s, \gamma_\lambda)\Big|_{\lambda=0} =& \delta_1 \int_0^{\bar{c}} \left(\frac{\partial}{\partial \lambda} \gamma_\lambda(c)\Big|_{\lambda=0} \right) \mathrm{d}c + \delta_2 \int_0^k \left(\frac{\partial}{\partial \lambda} \gamma_\lambda(c)\Big|_{\lambda=0} \right) \mathrm{d}c + \delta_3 \frac{\partial}{\partial \lambda} \gamma_\lambda(k)\Big|_{\lambda=0} \\
=& \delta_1 \xi_{\kappa^*}(\bar{c}|0) + \delta_2 \xi_{\kappa^*}(k|0) + \delta_3 \xi'_{\kappa^*}(k|0) \\
=& \kappa^*(k) \\
<& 0
\end{aligned}$$

and yet

$$\frac{\partial}{\partial \lambda} \left(\int_0^c \gamma_\lambda(x)\,\mathrm{d}x \right) \Big|_{\lambda=0} = \int_0^c \left(\frac{\partial}{\partial \lambda} \gamma_\lambda(x)\Big|_{\lambda=0} \right) \mathrm{d}x = \xi_{\kappa^*}(c|0) > 0,$$

for all $c \in [0, \bar{c}]$. As a result, within the family of (regular) technologies $\{\gamma_\lambda | \lambda \in [0, 1]\}$, for $\lambda > 0$ small enough, it must be that

$$\int_0^c [\gamma_\lambda(x) - \gamma(x)]\,\mathrm{d}x > 0,$$

for all $c \in [0, \bar{c}]$; and that

$$W^{\alpha,\beta}(s, \gamma) > W^{\alpha,\beta}(s, \gamma_\lambda).$$

This completes the proof. ■